

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

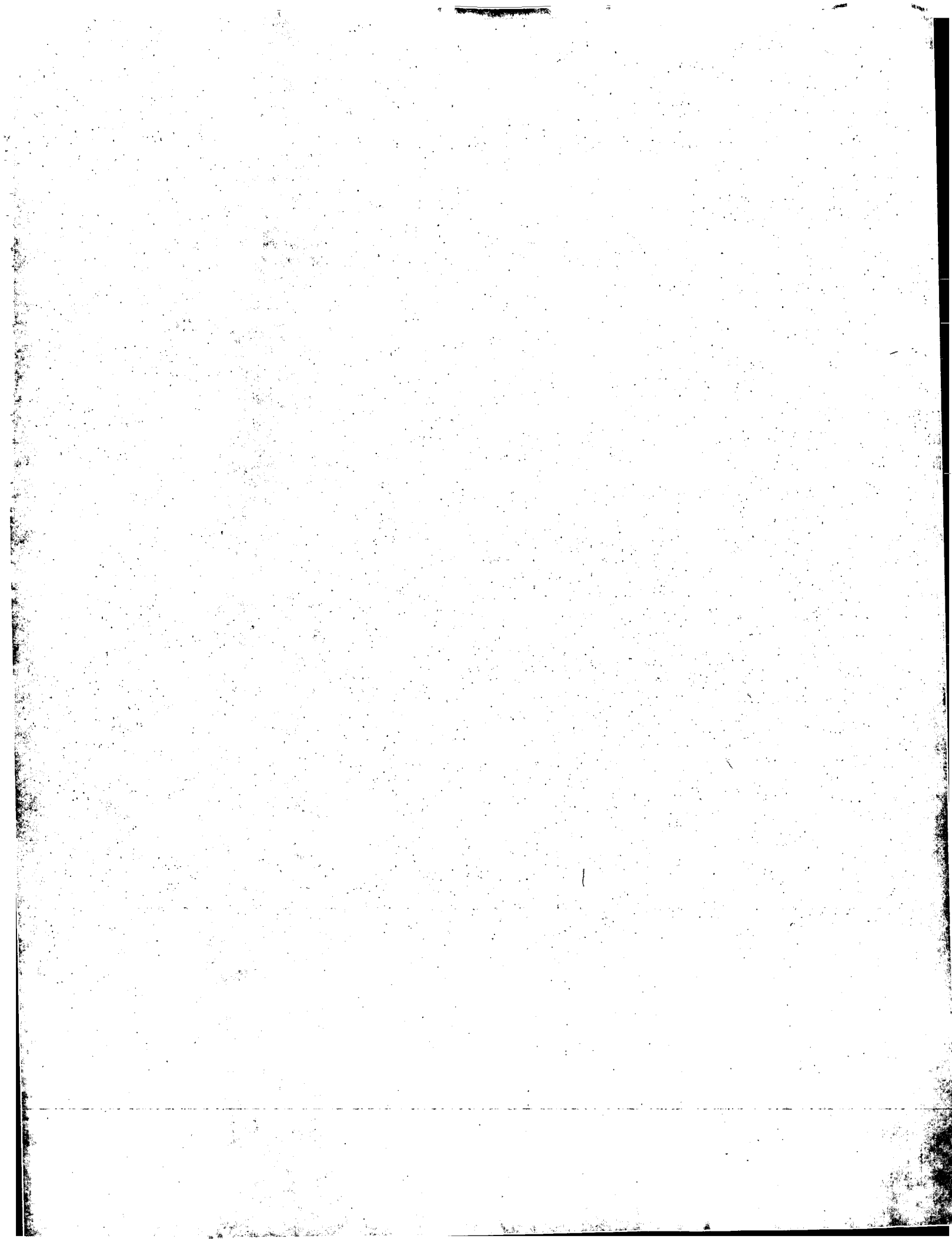
Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 01-17-03
Searcher: Beverly e4994
Terminal time: 26
Elapsed time: _____
CPU time: _____
Total time: 28
Number of Searches: 1
Number of Databases: _____

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG Suite
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 6615.05 seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: .1,accttgggtgctgggacg.....agagtttttgaattttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3984	100.0	3984	6	AR201455	AR201455 Sequence
2	3980.4	99.9	4307	9	AF104413	AF104413 Homo sapi
3	3775	94.8	4241	9	AF164041	AF164041 Homo sapi
4	2483	62.3	3213	6	AR201456	AR201456 Sequence
5	2483	62.3	3213	10	AF104414	AF104414 Mus muscu
6	2422	60.8	2442	6	E24613	E24613 warts Prote
7	1518.2	38.1	80146	9	AL583963	AL583963 Human DNA
8	1360.4	34.1	1374	6	E24614	E24614 warts Prote
9	930.8	23.4	1807	17	AF119846	AF119846 Homo sapi
10	848.4	21.3	16490	2	AC102712	AC102712 Mus muscu
11	788.6	19.8	795	6	E24616	E24616 warts Prote
12	760	19.1	3155	6	AR201457	AR201457 Sequence
13	760	19.1	3460	10	AB023958	AB023958 Mus muscu
14	744.4	18.7	5486	6	E38226	E38226 Human tumor
15	744.4	18.7	5486	6	E38227	E38227 Human tumor
16	742.8	18.6	4098	9	AF207547	AF207547 Homo sapi
17	741.8	18.6	3412	9	AB028019	AB028019 Homo sapi
18	680.4	17.1	164490	2	AC102712	AC102712 Mus muscu
19	604.8	15.2	191604	2	AC126128	AC126128 Rattus no
20	580.4	14.6	582	6	E24615	E24615 warts Prote
21	526.6	13.2	181750	2	AC098623	AC098623 Rattus no
22	503.2	12.6	181750	2	AC098623	AC098623 Rattus no
23	470.6	11.8	5360	3	DROWARTS	L39837 Drosophila
24	469	11.8	5720	3	DMU29608	U29608 Drosophila
25	469	11.8	5720	6	AR201454	AR201454 Sequence
26	340	8.5	183597	9	AL356285	AL356285 Human DNA
27	281.2	7.1	1322	3	EOC249683	AJ249683 Euplotes
28	281.2	7.1	1578	3	EOC249680	AJ249680 Euplotes
29	275	6.9	1804	3	EOC249682	AJ249682 Euplotes
30	269.4	6.8	1935	6	AR139101	AR139101 Sequence
31	269.4	6.8	4725	9	BC028603	BC028603 Homo sapi
32	269.4	6.8	4983	6	AX056364	AX056364 Sequence
33	269.4	6.8	5181	9	AB023182	AB023182 Homo sapi
34	269	6.8	2202	8	MCPRTKINA	Z30329 M.crySTALLI
35	255.6	6.4	2422	8	AY128279	AY128279 Arabidops
36	255.6	6.4	1606	3	EOC249681	AJ249681 Euplotes
37	255.4	6.4	1314	3	EOC249684	AJ249684 Euplotes
38	255.4	6.4	1745	3	AF275634	AF275634 Caenorhab
39	254.6	6.4	3018	6	A52140	A52140 Sequence 6
40	254.6	6.4	3018	6	AR084691	AR084691 Sequence
41	254.6	6.4	3018	9	HSPROKINX	Z35102 H.sapiens m
42	254.6	6.4	3593	9	BC012085	BC012085 Homo sapi
43	250.6	6.3	1738	8	AY096444	AY096444 Arabidops
44	250.6	6.3	2365	8	AY063969	AY063969 Arabidops
45	248.6	6.2	1807	3	ECU47679	U47679 Euplotes cr

ALIGNMENTS

RESULT 1	AR201455	3984 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201455	Sequence 3 from patent US 6359193.			
DEFINITION	AR201455	Sequence 3 from patent US 6359193.			
ACCESSION	AR201455	GI:20252343			
VERSION	AR201455.1				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 3984)				
AUTHORS	Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.				
TITLE	Nucleotide sequences of lats genes				
JOURNAL	Patent: US 6359193-A 3 19-MAR-2002;				
FEATURES	Location/Qualifiers				

source		1..3984	/organism="unknown"			
BASE COUNT		1280 a	847 c	798 g	1059 t	
ORIGIN						
Query Match		100.0%	Score 3984;	DB 6;	Length 3984;	
Best Local Similarity		100.0%	Pred. No. 0;			
Matches 3984;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	ACCTTTGGGTGCTGGGAGGACTCTGGCGCTCAGCTCAGCTCGCGCTCAGGCGCCCTGAGGCGCCGTGGCC	60			
DB	1	ACCTTTGGGTGCTGGGAGGACTCTGGCGCTCAGCTCAGCTCGCGCTCAGGCGCCCTGAGGCGCCGTGGCC	60			
QY	61	GCTGTCCAGAGGCTCTGCTCTCCCTCCAGAGTTAAATTTATATATGTAAGAATTTTA	120			
DB	61	GCTGTCCAGAGGCTCTGCTCTCCCTCCAGAGTTAAATTTATATATGTAAGAATTTTA	120			
QY	121	ACAGTCTCGGGGACTTCTTGAAGGATCATTTTCACTTTTGCTCAGAAGAAAGCTCTGGA	180			
DB	121	ACAGTCTCGGGGACTTCTTGAAGGATCATTTTCACTTTTGCTCAGAAGAAAGCTCTGGA	180			
QY	181	TCTATCAATAAAGAAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGA	240			
DB	181	TCTATCAATAAAGAAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGA	240			
QY	241	GTGAAAGCCAGAGGATATAGACAAATGAGGCTAAGACCTTTCCTGCCAGTAACATA	300			
DB	241	GTGAAAGCCAGAGGATATAGACAAATGAGGCTAAGACCTTTCCTGCCAGTAACATA	300			
QY	301	CTGTCACTAGCCGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC	360			
DB	301	CTGTCACTAGCCGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC	360			
QY	361	CATCTGATGCTCTAAGCTGAGCATAACATGAGTAATAATGTCAACGAGAGTCCCTGCAC	420			
DB	361	CATCTGATGCTCTAAGCTGAGCATAACATGAGTAATAATGTCAACGAGAGTCCCTGCAC	420			
QY	421	AAGTCAGAAATCCACCCAAATTTGGGAGCATATAAGGCTTCAGGAAATTCGAAACT	480			
DB	421	AAGTCAGAAATCCACCCAAATTTGGGAGCATATAAGGCTTCAGGAAATTCGAAACT	480			
QY	481	CTCTGCTTCCATTTGCAATGAACAAATTTCTCTCGAGTACTTCAGAGTAAATCCAC	540			
DB	481	CTCTGCTTCCATTTGCAATGAACAAATTTCTCTCGAGTACTTCAGAGTAAATCCAC	540			
QY	541	AAATGCTTCAGACTTGAAGCTGCTGATTTGATGAGGATATGTTATCAAGCTCTTC	600			
DB	541	AAATGCTTCAGACTTGAAGCTGCTGATTTGATGAGGATATGTTATCAAGCTCTTC	600			
QY	601	AGAAACTTAACAACAGAGTATAGAGCAGCAATGAAATTCATTAGTAAATGAGTTACC	660			
DB	601	AGAAACTTAACAACAGAGTATAGAGCAGCAATGAAATTCATTAGTAAATGAGTTACC	660			
QY	661	AAATGCTTCGAGCAGAGGATGCTGAGAGCTGCGAGAGCTTAAATGCCAGCATGA	720			
DB	661	AAATGCTTCGAGCAGAGGATGCTGAGAGCTGCGAGAGCTTAAATGCCAGCATGA	720			
QY	721	ACCAGGGAATGCGAGCAATCAGTTAACCGCAACAGAGCTGGAAGGTTCTTAAAGAA	780			
DB	721	ACCAGGGAATGCGAGCAATCAGTTAACCGCAACAGAGCTGGAAGGTTCTTAAAGAA	780			
QY	781	CCTTAGTCTCAGAGGCTGCGCGCCACTAGGAGAAAGTGTGCCCTATCATCTGAGA	840			
DB	781	CCTTAGTCTCAGAGGCTGCGCGCCACTAGGAGAAAGTGTGCCCTATCATCTGAGA	840			
QY	841	GTCCCAACTACAGACAGATGAGGAGACCTTGTCTGGATCTGATATACAGCATTTG	900			
DB	841	GTCCCAACTACAGACAGATGAGGAGACCTTGTCTGGATCTGATATACAGCATTTG	900			
QY	901	TTCAAGCTCACCTACGACGAGAGAGTGAACCCGCCACCCACCTCAAGTAAGA	960			
DB	901	TTCAAGCTCACCTACGACGAGAGAGTGAACCCGCCACCCACCTCAAGTAAGA	960			

Db 2041 ATAAAGAAAGAAACAGAGATTACAACTTCACCTATTACTGTTAGGAAAAACAAGAAAGATG 2100
Qy |||||||
Db 2101 AAGACGGAAGGAATCTCGTATTCAAAGTTATTCTCCTCAAGCAATTTAAATTTCTTTATGG 2160
Qy |||||||
Db 2101 AAGACGGAAGGAATCTCGTATTCAAAGTTATTCTCCTCAAGCAATTTAAATTTCTTTATGG 2160
Qy |||||||
Db 2161 AGCAACATGTAGAAAATGACTCAATCTCATCAGCAGCGCTACATCGTAAAAAACAAAT 2220
Qy |||||||
Db 2221 TAGAGAATGAATGATCGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGA 2280
Qy |||||||
Db 2221 TAGAGAATGAATGATCGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGA 2280
Qy |||||||
Db 2281 TGCTTTGCCAAAAAAGTCTAATTACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2340
Qy |||||||
Db 2281 TGCTTTGCCAAAAAAGTCTAATTACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2340
Qy |||||||
Db 2341 TGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAA 2400
Qy |||||||
Db 2341 TGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAA 2400
Qy |||||||
Db 2401 AAGTAGACTAAGGCTTTGTATGCAACAAAACTCTTCGAAAGAAAGATGTTCTTCCTC 2460
Qy |||||||
Db 2401 AAGTAGACTAAGGCTTTGTATGCAACAAAACTCTTCGAAAGAAAGATGTTCTTCCTC 2460
Qy |||||||
Db 2461 GAAATCAAGTCGCTCATGTTAAGCTGTAGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
Qy |||||||
Db 2461 GAAATCAAGTCGCTCATGTTAAGCTGTAGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
Qy |||||||
Db 2521 GGGTAGTTCGCTATATTATTCATTCCAAAGATAAGGACAAATTTATACCTTTGTAATGGACT 2580
Qy |||||||
Db 2521 GGGTAGTTCGCTATATTATTCATTCCAAAGATAAGGACAAATTTATACCTTTGTAATGGACT 2580
Qy |||||||
Db 2581 ACATTCCTGGGGGTGATGATGAGCCCTATTAAATAGAAATGGGCACTCTTCAGAAAGTC 2640
Qy |||||||
Db 2581 ACATTCCTGGGGGTGATGATGAGCCCTATTAAATAGAAATGGGCACTCTTCAGAAAGTC 2640
Qy |||||||
Db 2641 TGGCAGCAATCTACATAGCAAGCACTTACCTGTGCAAGTTGAAAGTCTTCAATAAATGGGTT 2700
Qy |||||||
Db 2641 TGGCAGCAATCTACATAGCAAGCACTTACCTGTGCAAGTTGAAAGTCTTCAATAAATGGGTT 2700
Qy |||||||
Db 2701 TTATTCATAGAGATATTAAACCTGATAATATTTTTCATGATCGTGTGATGATGATGATGAT 2760
Qy |||||||
Db 2701 TTATTCATAGAGATATTAAACCTGATAATATTTTTCATGATCGTGTGATGATGATGATGAT 2760
Qy |||||||
Db 2761 TGACTGACTTTGGCTCTGCACTGGCTTCAGATGACACACAGATCTTAAGTACTATCAGA 2820
Qy |||||||
Db 2761 TGACTGACTTTGGCTCTGCACTGGCTTCAGATGACACACAGATCTTAAGTACTATCAGA 2820
Qy |||||||
Db 2821 GTGGTGACCATCCCGCAAGATAGCATGGATTTCAGTAATGAATGGGGGATCCCTCAA 2880
Qy |||||||
Db 2821 GTGGTGACCATCCCGCAAGATAGCATGGATTTCAGTAATGAATGGGGGATCCCTCAA 2880
Qy |||||||
Db 2881 GCTGTGATGTGGACAGACTGAAGCCATTAGAGCGGAGAGCTGCAGCCAGCAGCAGCAGC 2940
Qy |||||||
Db 2881 GCTGTGATGTGGACAGACTGAAGCCATTAGAGCGGAGAGCTGCAGCCAGCAGCAGCAGC 2940
Qy |||||||
Db 2941 GATGTCTAGACATCTTTGGTTGGGACTCCCAATTAATATGACCTCAAGTGTGCTAC 3000
Qy |||||||
Db 2941 GATGTCTAGACATCTTTGGTTGGGACTCCCAATTAATATGACCTCAAGTGTGCTAC 3000
Qy |||||||
Db 3001 GAACAGGATACACAGATGTTGATGTTGGTGGAGTGGTGTATTTCTTTTGAATGCT 3060
Qy |||||||
Db 3001 GAACAGGATACACAGATGTTGATGTTGGTGGAGTGGTGTATTTCTTTTGAATGCT 3060
Qy |||||||
Db 3061 TGGTGGGCAACCTCTTTCTTGGCACAACACACCACTTGAACACACAAATGAAGTTATCA 3120
Qy |||||||
Db 3061 TGGTGGGCAACCTCTTTCTTGGCACAACACACCACTTGAACACACAAATGAAGTTATCA 3120
Qy |||||||
Db 3121 ACTGCAACATCTCTTCCATTTCCACCAAGCTTAAGTCTAGTCTCTGAGCTCTGATC 3180
Qy |||||||

Db 3121 ACTGCAACATCTCTTCCATTTCCACATTCACCACCAAGCTAAACTCAGTCTGAGCTTCTGATC 3180
Qy |||||||
Db 3181 TTATTATTAAACTTTTGGCGAGGACCGAAGATCGCTTAGCGAAGATGCTGCTGATGAA 3240
Qy |||||||
Db 3181 TTATTATTAAACTTTTGGCGAGGACCGAAGATCGCTTAGCGAAGATGCTGCTGATGAA 3240
Qy |||||||
Db 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTTCCAGTGCACCTGAGACAGCAGCTG 3300
Qy |||||||
Db 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTTCCAGTGCACCTGAGACAGCAGCTG 3300
Qy |||||||
Db 3301 CTTTCATACATTTCTTAAATCACACACCCCAAGATACATCAATTTTGTGCTGTTGATC 3360
Qy |||||||
Db 3301 CTTTCATACATTTCTTAAATCACACACCCCAAGATACATCAATTTTGTGCTGTTGATC 3360
Qy |||||||
Db 3361 CTGATAAATTTATGGAGTGATGATAAGGAGAGAAATGTAATGACACTCTCAATGGAT 3420
Qy |||||||
Db 3361 CTGATAAATTTATGGAGTGATGATAAGGAGAGAAATGTAATGACACTCTCAATGGAT 3420
Qy |||||||
Db 3421 GGTATAAATTTGAAGCATCTCTGAACATTCATTTCTTATGAATTTACCTTCCGAAGTCTT 3480
Qy |||||||
Db 3421 GGTATAAATTTGAAGCATCTCTGAACATTCATTTCTTATGAATTTACCTTCCGAAGTCTT 3480
Qy |||||||
Db 3481 TTGATGACAAATGGCTTACCCATATAATTTCCGAAGCTTATGAAATGATGATGATGATGAT 3540
Qy |||||||
Db 3481 TTGATGACAAATGGCTTACCCATATAATTTCCGAAGCTTATGAAATGATGATGATGATGAT 3540
Qy |||||||
Db 3541 CACAAGGCTCAGACAGCAGCTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAANA 3600
Qy |||||||
Db 3541 CACAAGGCTCAGACAGCAGCTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAANA 3600
Qy |||||||
Db 3601 ATCGCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAANAAG 3660
Qy |||||||
Db 3601 ATCGCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAANAAG 3660
Qy |||||||
Db 3661 GGCCTGAATGCGAGCTGTTTTCAGGCTCTGAGAGTAAATTAATGCAATATGACAGAGC 3720
Qy |||||||
Db 3661 GGCCTGAATGCGAGCTGTTTTCAGGCTCTGAGAGTAAATTAATGCAATATGACAGAGC 3720
Qy |||||||
Db 3721 TATATGCTGCTCTGTGTACAAATTTTATTTTCCCTAAATTTATGGGAAATCCCTTTAA 3780
Qy |||||||
Db 3721 TATATGCTGCTCTGTGTACAAATTTTATTTTCCCTAAATTTATGGGAAATCCCTTTAA 3780
Qy |||||||
Db 3781 AATGTTAATTTATCCAGCCGTTTAAATCAGTATTTAGAAAAAATTTGTTAAGGAAAG 3840
Qy |||||||
Db 3781 AATGTTAATTTATCCAGCCGTTTAAATCAGTATTTAGAAAAAATTTGTTAAGGAAAG 3840
Qy |||||||
Db 3841 TAAATTTGAACCTGAATTTATATAGTCACTTCTGTTTAAAGTAAAGTAAAGTAG 3900
Qy |||||||
Db 3841 TAAATTTGAACCTGAATTTATATAGTCACTTCTGTTTAAAGTAAAGTAAAGTAG 3900
Qy |||||||
Db 3901 TGCTTTGTTTAAAGGAGAAACCTGCTATCTATTTTGTATATATGCTAAATAATTTTAA 3960
Qy |||||||
Db 3901 TGCTTTGTTTAAAGGAGAAACCTGCTATCTATTTTGTATATATGCTAAATAATTTTAA 3960
Qy |||||||
Db 3961 TACAAGAGTTTGTGAATTTTTT 3984
Qy |||||||
Db 3961 TACAAGAGTTTGTGAATTTTTT 3984

RESULT 2

AF104413

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF104413 Homo sapiens large tumor suppressor 1 (LATS1) mRNA, complete cds.
4307 bp linear

AF104413.1 GI:4324433

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4307)

Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.

Identifying tumor suppressors in genetic mosaics: the Drosophila

QY 1203 GTTGCAGACAAACCAATCATCATGAGAGTCTTAGCAAAATTTAACTTTCCATCAGGGAGA 1262
DB 1521 GTTGCAGACAAACCAATCATCATGAGAGTCTTAGCAAAATTTAACTTTCCATCAGGGAGA 1580
QY 1263 CCTGGAATGCGAATGTTGCTAGGAACTGATTTATTCATGATACACAAATTTGTCCT 1322
DB 1581 CTTGGAATGCGAATGTTGCTAGGAACTGATTTATTCATGATACACAAATTTGTCCT 1640
QY 1323 GCTGCACTGTGAATGCGCAGGACCACTCCATATCTCTGACAGCAGCTAATGACAA 1382
DB 1641 GCTGCACTGTGAATGCGCAGGACCACTCCATATCTCTGACAGCAGCTAATGACAA 1700
QY 1383 AGCCCTTCTGCTTTTACAAACAGGGGATGCTGCTCTCTCTCATATACAAATGGAAGT 1442
DB 1701 AGCCCTTCTGCTTTTACAAACAGGGGATGCTGCTCTCTCATATACAAATGGAAGT 1760
QY 1443 ATTCTCTAGTCTATGATGGTGCACAGAAATAGTCTATACATGAAGCTATATACAT 1502
DB 1761 ATTCTCTAGTCTATGATGGTGCACAGAAATAGTCTATACATGAAGCTATATACAT 1820
QY 1503 AGTGTACCTGGACTGCAACAAATGGCTCAGTCTCTCTGCTCCAGCCAGTCATCC 1562
DB 1821 AGTGTACCTGGACTGCAACAAATGGCTCAGTCTCTCTGCTCCAGCCAGTCATCC 1880
QY 1563 CCGAGCAGTGGGCATGAAATCCCTACATGGCAACCTAACATACCACTGAGGTCAAAATCT 1622
DB 1881 CCGAGCAGTGGGCATGAAATCCCTACATGGCAACCTAACATACCACTGAGGTCAAAATCT 1940
QY 1623 TTTAATAACCCATAGGAATAGCAAGTCTACTCTGCTAATTTTCAGCCTTCTGCTACA 1682
DB 1941 TTTAATAACCCATAGGAATAGCAAGTCTACTCTGCTAATTTTCAGCCTTCTGCTACA 2000
QY 1683 ACAGTCACTGCAATTTACACCACTGCTTATCAACAGCCTGTGAAAGTATGCTGTATTA 1742
DB 2001 ACAGTCACTGCAATTTACACCACTGCTTATCAACAGCCTGTGAAAGTATGCTGTATTA 2060
QY 1743 AAACAGAGCTACAGACTGCTTTAGCACCTACACACCCTCTTTGGATACCAAGCAAT 1802
DB 2061 AAACAGAGCTACAGACTGCTTTAGCACCTACACACCCTCTTTGGATACCAAGCAAT 2120
QY 1803 CAACTGTTCAACCACTGCTTTTCTGAGGAAACCGCTTCAATGTGACTGTGATGCA 1862
DB 2121 CAACTGTTCAACCACTGCTTTTCTGAGGAAACCGCTTCAATGTGACTGTGATGCA 2180
QY 1863 CTTGTTGCTGAAGCTTCAACCACTATCAAGGACCAACCACTTACCACTTACCACTTACCACT 1922
DB 2181 CTTGTTGCTGAAGCTTCAACCACTATCAAGGACCAACCACTTACCACTTACCACTTACCACT 2240
QY 1923 CACCAAAACCCATCTGTTCTCCATACAGTCAATCAGTAAGCCTTAGCAAGAGGATCAG 1982
DB 2241 CACCAAAACCCATCTGTTCTCCATACAGTCAATCAGTAAGCCTTAGCAAGAGGATCAG 2300
QY 1983 CCAAGCTTGCACCAAGAGATGAGAGTGAAGAGATTTAGAAATGTTAGTAGTGGGAT 2042
DB 2301 CCAAGCTTGCACCAAGAGATGAGAGTGAAGAGATTTAGAAATGTTAGTAGTGGGAT 2360
QY 2043 AAAGAAAAGAACAGATTACACTTCACTTATCTGTTAGGAAAACAAAGAAAGATGAA 2102
DB 2361 AAAGAAAAGAACAGATTACACTTCACTTATCTGTTAGGAAAACAAAGAAAGATGAA 2420
QY 2103 GAGCGAAGGGAATCTGATTCAAAAGTTATTTCTCTCAAGCATTTTAAATTTCTTTATGGAG 2162
DB 2421 GAGCGAAGGGAATCTGATTCAAAAGTTATTTCTCTCAAGCATTTTAAATTTCTTTATGGAG 2480
QY 2163 CAACATGTAGAAATGTACTCAAAATCTCATCAGCAGCCTTACATCGTAAAAACAATTA 2222
DB 2481 CAACATGTAGAAATGTACTCAAAATCTCATCAGCAGCCTTACATCGTAAAAACAATTA 2540
QY 2223 GAGAAATGATCGCGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATG 2282
DB 2541 GAGAAATGATCGCGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATG 2600

QY 2283 CTTTGCCAAAAGAAATTAATTTACATCCGTCTTAAAGGGCTTAAATGCAAGTCTATG 2342
DB 2601 CTTTGCCAAAAGAAATTAATTTACATCCGTCTTAAAGGGCTTAAATGCAAGTCTATG 2660
QY 2343 TTTGTGAAGATAAAGACACTAGGAATAGAGCATTTGGTGAAGTCTGTCTAGCAAGAAA 2402
DB 2661 TTTGTGAAGATAAAGACACTAGGAATAGAGCATTTGGTGAAGTCTGTCTAGCAAGAAA 2720
QY 2403 GTAGATACCTAAGCTTTGTATGCAACAAACTCTTCCAAAGAAAGATGTTCTTCTCGA 2462
DB 2721 GTAGATACCTAAGCTTTGTATGCAACAAACTCTTCCAAAGAAAGATGTTCTTCTCGA 2780
QY 2463 AATCAAGTCGCTCATCTTAAAGCTGAGAGATATCTCGGCTGAAGCTGCAAAATGAATGG 2522
DB 2781 AATCAAGTCGCTCATCTTAAAGCTGAGAGATATCTCGGCTGAAGCTGCAAAATGAATGG 2840
QY 2523 GTAGTTCGCTCATATTTATTCATTCACAGATAAGGACAATTTATCTTTGTATGAGTAC 2582
DB 2841 GTAGTTCGCTCATATTTATTCATTCACAGATAAGGACAATTTATCTTTGTATGAGTAC 2900
QY 2583 ATTCTCTGGGGTGATGATGAGCCTTATTAATGAAATGGGCTCTTTCCAGAAAGTCTG 2642
DB 2901 ATTCTCTGGGGTGATGATGAGCCTTATTAATGAAATGGGCTCTTTCCAGAAAGTCTG 2960
QY 2643 GCACGATTTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTT 2702
DB 2961 GCACGATTTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTT 3020
QY 2703 ATTCTAGAGATATTAACCTGATTAATTTTGTATGATGCTGATGCTCATATTAATTTG 2762
DB 3021 ATTCTAGAGATATTAACCTGATTAATTTTGTATGATGCTGATGCTCATATTAATTTG 3080
QY 2763 ACTGACTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2822
DB 3081 ACTGACTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3140
QY 2823 GGTGACCATCCACGCAAGATAGATGATTTTCAATTAATGAGGGGGATCCCTCAAGC 2882
DB 3141 GGTGACCATCCACGCAAGATAGATGATTTTCAATTAATGAGGGGGATCCCTCAAGC 3200
QY 2883 TGTCCATGTGAGACAGACTGAGCCNTTAGAGCGGAGCTGACGCGCAGCAGCAGGCA 2942
DB 3201 TGTCCATGTGAGACAGACTGAGCCNTTAGAGCGGAGCTGACGCGCAGCAGCAGGCA 3260
QY 2943 TGTCTAGACACATCTTTGTTGGGACTCCCAATTAATTTGACCTGAAGTGTGCTACGA 3002
DB 3261 TGTCTAGACACATCTTTGTTGGGACTCCCAATTAATTTGACCTGAAGTGTGCTACGA 3320
QY 3003 ACAGGATACACAGCTTTGTTGTTGGGAGTGTGTTGTTTATTTCTTTTGAATGTTG 3062
DB 3321 ACAGGATACACAGCTTTGTTGTTGGGAGTGTGTTGTTTATTTCTTTTGAATGTTG 3380
QY 3063 GTGGGACAACCTCTTTCTTGGCACAACACCATTTAGAAACACAAATGAAGTTFATCAAC 3122
DB 3381 GTGGGACAACCTCTTTCTTGGCACAACACCATTTAGAAACACAAATGAAGTTFATCAAC 3440
QY 3123 TGGCAAAACATCTCTTCAATTTCCACCAAGCTAAGCTCAGTCCGAGCTTCTGATCTT 3182
DB 3441 TGGCAAAACATCTCTTCAATTTCCACCAAGCTAAGCTCAGTCCGAGCTTCTGATCTT 3500
QY 3183 ATTATTAACCTTTGCGGAGGACCCGAAGATCGCTTTAGCAAGAATGGTCTGATGAAATA 3242
DB 3501 ATTATTAACCTTTGCGGAGGACCCGAAGATCGCTTTAGCAAGAATGGTCTGATGAAATA 3560
QY 3243 AAAGCTCATCATTTTAAACAAATTTGCTTCTCCAGTACCTGAGCAGCAGTCTGCT 3302
DB 3561 AAAGCTCATCATTTTAAACAAATTTGCTTCTCCAGTACCTGAGCAGCAGTCTGCT 3620
QY 3303 TCATACATTTCTTAAATCAGACCCCAACAGATACATCAATTTTGTATCTTCTGATCTT 3362
DB 3621 TCATACATTTCTTAAATCAGACCCCAACAGATACATCAATTTTGTATCTTCTGATCTT 3680
QY 3363 GATAAATTTAGGATGATGATAACGAGGAAGAAATGTAATGACACTCTCTCAATGGATGG 3422

Fri Jan 17 11:16:52 2003

```

|||||
Db 3681 GATAAATATGGAGTATGATATAACGAGGAGAAATGTAATGACACTCTCAATGATGG 3740
Qy 3423 TATATAAATGGAAGCATCTGACATGATCTATGAAATTTACCTTCGGAAGTTTTTT 3482
Db 3741 TATAAATGGAAGCATCTGACATGATCTATGAAATTTACCTTCGGAAGTTTTTT 3800
Qy 3483 GATGACAATGGCTACCATATAATATCCGAAGCCTATGAAATGATGAAATATCA 3542
Db 3801 GATGACAATGGCTACCATATAATATCCGAAGCCTATGAAATGATGAAATATCA 3860
Qy 3543 CAAGGCTCAGACGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3602
Db 3861 CAAGGCTCAGACGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 3920
Qy 3603 CGCATCTAGTATATGTTTAAACACACTAGTAAATGATGATGATGATGATGATGATG 3662
Db 3921 CGCATCTAGTATATGTTTAAACACACTAGTAAATGATGATGATGATGATGATGATG 3980
Qy 3663 CCTGAAATGCGAGGTGTTTGAAGTCTGAGAGTAAATGATGATGATGATGATGATGATG 3722
Db 3981 CCTGAAATGCGAGGTGTTTGAAGTCTGAGAGTAAATGATGATGATGATGATGATGATG 4040
Qy 3723 TATATGTCCTCTGTCGATATATTTTCTTAAATGATGATGATGATGATGATGATGATG 3782
Db 4041 TATATGTCCTCTGTCGATATATTTTCTTAAATGATGATGATGATGATGATGATGATG 4100
Qy 3783 TGTTAAATTTATCCAGCGCTTAAATGATGATGATGATGATGATGATGATGATGATGATG 3842
Db 4101 TGTTAAATTTATCCAGCGCTTAAATGATGATGATGATGATGATGATGATGATGATGATG 4160
Qy 3843 AATTATGAATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3902
Db 4161 AATTATGAATGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4220
Qy 3903 CTTTCTTAAAGGAGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 3962
Db 4221 CTTTCTTAAAGGAGAAACCTGATGATGATGATGATGATGATGATGATGATGATGATG 4280
Qy 3963 CAAGAGTTTTGAAATTTTTTT 3984
Db 4281 CAAGAGTTTTGAAATTTTTTT 4302

```

```

RESULT 3
AF164041 4241 bp mRNA linear PRI 06-DEC-1999
LOCUS
DEFINITION Homo sapiens WARTS protein kinase (WARTS) mRNA, complete cds.
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 4241)
Kishimoto,T., Niva,S.-i., Nagamine,Y., Nishiyama,Y. and Saya,H.
WARTS protein, polynucleotide encoding the same, antisense
polynucleotide thereof, and antibody recognizing the protein
Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-ku,
Yokohama-shi; Kanagawa;
Japan;
2. (bases 1 to 4241)
Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiroka,T., Kitamura,N. and
Saya,H.
A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
FEBS Lett. 459 (2), 159-165 (1999)
JOURNAL 99447636
MEDLINE
PUBMED 10518011

```

```

REFERENCE
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto.860-0811,
Japan

```

```

FEATURES
source
1. 4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 4241
/gene="WARTS"
33..3425
/gene="WARTS"

```

```

note="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"
/db_xref="GI:5738136"
/translation="MKRSEKPEGYPMRPKTPFASNYTVSSROMLOEIRSLNLSKP
SDAAKHNMSKSTEDPQVNPFPKFTGHKALQEIENSLPEANETNSRSTSEVN
PQMLQDQAAGFEDMVIQALQKTNNSIERAIEFISKMSQDPRREQMAAARPIN
ASMKPGNVQSNRRKQSWKSKESLVQPHGPPGSGVAYHSESPNSQTDVGRPLSGS
GISAFVQAHPSNGRVNPPPPQVRSVTPPPRGTTPPPPTPPPSWEPNSQTKR
YSGNMEYVLSRTPVPGAWOEGYPPPLNTSPMPPNQGGISSLVPGQSPALQMG
SKNFPSPGRGCMQGTGOTDFMIHONVVPAGTVNRPVPLTAANGQSPALQMG
GSAAPSSYTGSIQSMVMVNRNSHNMELNLSVPGLOTNMPQSSAPASQSSPSGHE
IPTQNPPIVRNSENNPNRNSHNSANSQPSATTVTATITAPIQPVKSMRLKPEL
QTALAPTHPSWIPIQTVQSPFPFEGTASNVMPVAENIPQPPPPPKHLLHQ
NPSPVPEISKSPKEDQSLPKEDSEKSEYENDSDGKKEKQITSPITVKNKKKE
ERRESIQSYSPQAFKFEHVENLVKSHQORLHRKQLENMRVGLSQQDAQDMR
KMLCKESNVIRLRAKMDKSMFKIKTLGAFGEVCLARKVDKLYATKTLRKDK
VILRQNVHVAERDILAEADNEWVRLYVSFQDKNLYFVMDYIPGDDMMSLLRMG
IPESLAREYIAELCAVESVHKMGFIHERDKPDLILDRQGHILKIDFGCLGTGRTP
NTIAYVSGDHPQSDMSFNSWGVILFEMLVGQPPFLAGTPLETOMKYNVQTSHP
POAKLSPEASDLIIKICRGPEDLKNGADEIKAHFFKTIIDFSSLRQOOSASYIPKI
THPTDTSNFDVPDPDKLWSDNEENNDLNGWYKNGKHPEHAFYEFFRFRFFDDING
YPNYPKPIEYEVINSQSEQSDDEDDQNTGSEIKNRDLVYV"

```

```

misc_feature
2145..3059
/gene="WARTS"
/notes="Region: serine/threonine kinase domain"
BASE COUNT 1384 a 884 c 817 g 1156 t
ORIGIN

```

```

Query Match 94.8% Score 3775; DB 9; Length 4241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3786; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

```

Qy 199 CCTTCGTGTCGGCTACATATATAGATGTTTTCATGAGAGAGAGTGAAAGCCAGAGGAT 258
Db 1 CCTTCGTGTCGGCTACATATATAGATGTTTTCATGAGAGAGAGTGAAAGCCAGAGGAT 60
Qy 259 ATAGACAATAGGCGCTAAGACCTTTCTCCAGTACTACTCTCAGTAGCGGCAAA 318
Db 61 ATAGACAATAGGCGCTAAGACCTTTCTCCAGTACTACTCTCAGTAGCGGCAAA 120
Qy 319 TGTTCACAGAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGCTGCTAAGG 378
Db 121 TGTTCACAGAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGCTGCTAAGG 180
Qy 379 CTGAGCATAACATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCAAGAAATCCACCCA 438
Db 181 CTGAGCATAACATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCAAGAAATCCACCCA 240
Qy 439 RAATTTGGGACGCATCATAAAGCCTTTCAGGAAATTCGAAACTCTCTGCTCCATTTGCAA 498
Db 241 RAATTTGGGACGCATCATAAAGCCTTTCAGGAAATTCGAAACTCTCTGCTCCATTTGCAA 300
Qy 499 ATGAAACAAATTTCTTCTCGGAGTACTTTCAGAGTTAATCCACAATGCTTCAAGACTTGC 558

```

Db	301	ATGAACAAATTCCTTCGGAGTACTTCAGAACTTAATCCCAAAATGCTTCAAGACTTGC	360
Qy	559	AAGTCGCTGGATTGTATGAGGATATGGTTATACAAGCTCTTCAGAAAACTAACACAGAA	618
Db	361	AAGTCGCTGGATTGTATGAGGATATGGTTATACAAGCTCTTCAGAAAACTAACACAGAA	420
Qy	619	GTATAGAAGCAGCAATGTAAATTCATTAGTAAATAGATTACCAAGATCCTCGACGAGAC	678
Db	421	GTATAGAAGCAGCAATGTAAATTCATTAGTAAATAGATTACCAAGATCCTCGACGAGAC	480
Qy	679	AGATGGCTGCAGCAGCTGCCAGACCTTAATATCCAGCATGAACACAGGAAATGTGCAGC	738
Db	481	AGATGGCTGCAGCAGCTGCCAGACCTTAATATCCAGCATGAACACAGGAAATGTGCAGC	540
Qy	739	AATCAGTTAAACCCCAACACAGAGCTGGAAAGGTTCTAAAGAAATCCTTAGTTCCTCAGAGC	798
Db	541	AATCAGTTAAACCCCAACACAGAGCTGGAAAGGTTCTAAAGAAATCCTTAGTTCCTCAGAGC	600
Qy	799	ATGCCCGCCACCTAGAGAAAGTGTGGCCCTATCATCTCGAGAGTCCCAACTCACAGACAG	858
Db	601	ATGCCCGCCACCTAGAGAAAGTGTGGCCCTATCATCTCGAGAGTCCCAACTCACAGACAG	660
Qy	859	ATGTAGGAAGACCTTGTCTGGATCTGGTATATACGACATTTGTTCAGCTCACCTAGCA	918
Db	561	ATGTAGGAAGACCTTGTCTGGATCTGGTATATACGACATTTGTTCAGCTCACCTAGCA	720
Qy	919	ACGGACAGAGAGTGAACCCCCACCAACACCTCAAGTAAGAGTGTACTCCTCCACCAC	978
Db	721	ACGGACAGAGAGTGAACCCCCACCAACACCTCAAGTAAGAGTGTACTCCTCCACCAC	780
Qy	979	CTCCAAGAGGCCAGACTCCCCCTCCAAGAGGTACAACTCCACCTCCCTTCATGGGAAC	1038
Db	781	CTCCAAGAGGCCAGACTCCCCCTCCAAGAGGTACAACTCCACCTCCCTTCATGGGAAC	840
Qy	1039	CAAACTCTCAACAAACGCGCTATTCGGAACATGGAATACCTATATCTCCGGAATCTCTC	1098
Db	841	CAAACTCTCAACAAACGCGCTATTCGGAACATGGAATACCTATATCTCCGGAATCTCTC	900
Qy	1099	CTGTCCCACCTGGGGCATGGCAAGGGGTATCTCTCCACCACTCTCAACACTTCTCCCCCA	1158
Db	901	CTGTCCCACCTGGGGCATGGCAAGGGGTATCTCTCCACCACTCTCAACACTTCTCCCCCA	960
Qy	1159	TGAATCTCTTAATCAAGGACAGAGAGGCATTAGTTCTGTTCCTGTGTGGCAGACAACAA	1218
Db	961	TGAATCTCTTAATCAAGGACAGAGAGGCATTAGTTCTGTTCCTGTGTGGCAGACAACAA	1020
Qy	1219	TCATCATGCAGAGTCTAGCAAAATTTAACTTTCCATCAGGGAGACCTTGGAAATGCAGATG	1278
Db	1021	TCATCATGCAGAGTCTAGCAAAATTTAACTTTCCATCAGGGAGACCTTGGAAATGCAGATG	1080
Qy	1279	GTACTGGACAAACTGATTTTCATGATACCAAAATGTTGTCCCTCTGGCACTGTGAATC	1338
Db	1081	GTACTGGACAAACTGATTTTCATGATACCAAAATGTTGTCCCTCTGGCACTGTGAATC	1140
Qy	1339	GGCAGCCACCACTCCATATTCCTCTGACAGCAGCTAATGGACAAAGCCCTTCTGCTTTAC	1398
Db	1141	GGCAGCCACCACTCCATATTCCTCTGACAGCAGCTAATGGACAAAGCCCTTCTGCTTTAC	1200
Qy	1399	AAACAGGGGATCTGCTGCTCTTCCGTATATACAAATGGAAGTATTCCTCAGTCTATGA	1458
Db	1201	AAACAGGGGATCTGCTGCTCTTCCGTATATACAAATGGAAGTATTCCTCAGTCTATGA	1260
Qy	1459	TGTTGCCAAACAGAAATAGTCAATACATGGAACATATAAACAATAGTCTACCTGGACTGC	1518
Db	1261	TGTTGCCAAACAGAAATAGTCAATACATGGAACATATAAACAATAGTCTACCTGGACTGC	1320
Qy	1519	AAACAAATGGCCCTCAGTCTATCTTCTGCCAGCCAGTCAATCCCGAGCAGTGGGCATG	1578
Db	1321	AAACAAATGGCCCTCAGTCTATCTTCTGCCAGCCAGTCAATCCCGAGCAGTGGGCATG	1380
Qy	1579	AAATPCCCTACATGGCAACCTTAACATACCAAGTGAAGTCAAAATTCCTTTTAATAAACCCATTAAG	1638

Db	1381	AAATCCCTACATGGCAACCTCAACATACCAGTGAGGTCAAATTTCTTTTAATAACCCATTAG	1414
Qy	1639	GAATAGAGCAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCTACAAACAGTCACTGCAATTA	1698
Db	1441	GAATAGAGCAAGTCACTCTGCTAAATTTCTAGCCTTCTGCTACACAGTCACTGCAATTA	1500
Qy	1699	CACAGCTCCTATTCAACAGCCTGTGAAAGTAGTGGTGATTTAAACACAGAGCTACAGA	1758
Db	1501	CACAGCTCCTATTCAACAGCCTGTGAAAGTAGTGGTGATTTAAACACAGAGCTACAGA	1560
Qy	1759	GTGCTTTAGCACCTACACACCCCTTCTTGATACACAGCCCAATTTCAAACCTGTTCAACCCA	1818
Db	1561	GTGCTTTAGCACCTACACACCCCTTCTTGATACACAGCCCAATTTCAAACCTGTTCAACCCA	1620
Qy	1819	GTCCCTTTTCTGAGGAAACCGCTTCAAATGTGACTGTGATGACACCTGTGCTGAAGCTC	1878
Db	1621	GTCCCTTTTCTGAGGAAACCGCTTCAAATGTGACTGTGATGACACCTGTGCTGAAGCTC	1680
Qy	1879	CAAACTATCAAGGACCAACCAACCCCTACCCAAACATCTGCTGACCAAAACCCATCTG	1938
Db	1681	CAAACTATCAAGGACCAACCAACCCCTACCCAAACATCTGCTGACCAAAACCCATCTG	1740
Qy	1939	TTCTCTCATACGAGTCAATCAGTTAAGCCTTAGCAAGAGGATCAGCCAAGCTTGCCCAAG	1998
Db	1741	TTCTCTCATACGAGTCAATCAGTTAAGCCTTAGCAAGAGGATCAGCCAAGCTTGCCCAAG	1800
Qy	1999	AAGATGAGAGTGAAAAGAGTTATGAAAATGTTGATAGTGGGATTAAGAAAAGAAACAGA	2058
Db	1801	AAGATGAGAGTGAAAAGAGTTATGAAAATGTTGATAGTGGGATTAAGAAAAGAAACAGA	1860
Qy	2059	TTACAATTCACCTATTACTGTGTAGAAAACAAAGAAAGATGAAGAGCGAAGGAATCTC	2118
Db	1861	TTACAATTCACCTATTACTGTGTAGAAAACAAAGAAAGATGAAGAGCGAAGGAATCTC	1920
Qy	2119	GTATTCAAAGTTATCTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAATG	2178
Db	1921	GTATTCAAAGTTATCTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAATG	1980
Qy	2179	TACTCAANTCTCATCAGCAGCGTCACTACATCGTAAAAACAAATTAGAGAAATCAAAATGATGC	2238
Db	1981	TACTCAANTCTCATCAGCAGCGTCACTACATCGTAAAAACAAATTAGAGAAATCAAAATGATGC	2040
Qy	2239	GGGTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAGAAAT	2298
Db	2041	GGGTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAGAAAT	2100
Qy	2299	CTAATTACATCCGCTTTAAAGGGCTAAATGGACAAGTCTATGTTGTGAAGATAAGA	2358
Db	2101	CTAATTACATCCGCTTTAAAGGGCTTTAAAGGGCTTTAAATGGACAAGTCTATGTTGTGAAGATAAGA	2160
Qy	2359	CACGTAGAAATAGGAGCATTTGGTGAAGTCTGCTAGCAAGAAAAGTACATACTAAGGCTT	2418
Db	2161	CACGTAGAAATAGGAGCATTTGGTGAAGTCTGCTAGCAAGAAAAGTACATACTAAGGCTT	2220
Qy	2419	TGATGCAACAAAAAATCTTTGCAAGAAAAGATGTTCTTCTTCGAAATCAAGTCGCTCATG	2478
Db	2221	TGATGCAACAAAAAATCTTTGCAAGAAAAGATGTTCTTCTTCGAAATCAAGTCGCTCATG	2280
Qy	2479	TTAAGGCTGAGAGATATCCTGGCTGAAGCTGACAAATGAATGGGTAGTTGCTCTATATT	2538
Db	2281	TTAAGGCTGAGAGATATCCTGGCTGAAGCTGACAAATGAATGGGTAGTTGCTCTATATT	2340
Qy	2539	ATTCAATTCGAAGATAAGGCAATTTATCTTTGTAATGSACTACATCTCCTGGGGTGATA	2598
Db	2341	ATTCAATTCGAAGATAAGGCAATTTATCTTTGTAATGSACTACATCTCCTGGGGTGATA	2400
Qy	2599	TGATGAGCCTATTAAATTAGAATGGGCATCTTTCCAGAAAGTCTGGCAGCATTTCTACATAG	2658
Db	2401	TGATGAGCCTATTAAATTAGAATGGGCATCTTTCCAGAAAGTCTGGCAGCATTTCTACATAG	2460
Qy	2659	CAGAACTTACCTGTGCAGTTGAAGTGTTCAATAAAATGGGTTTTATTCAATAGAGATATTA	2718
Db	2461	CAGAACTTACCTGTGCAGTTGAAGTGTTCAATAAAATGGGTTTTATTCAATAGAGATATTA	2520

QY 2719 AACCTGATAATATTTTGAATGATCGTGATGGTCAATATTAATTTGACTGACTTTGGCCTCT 2778
DB 2521 AACCTGATAATATTTTGAATGATCGTGATGGTCAATATTAATTTGACTGACTTTGGCCTCT 2580
QY 2779 GCACTGGCTTCAGATGGACACAGATTTCAAGTACTATCAGAGTGGTGGACCATCCACGGC 2838
DB 2581 GCACTGGCTTCAGATGGACACAGATTTCAAGTACTATCAGAGTGGTGGACCATCCACGGC 2640
QY 2839 AAGATACATGGATTTTCAAGTAAATGAATGGGGGATCCCTCAAGCTGTCTGATGTGGAGACA 2898
DB 2641 AAGATACATGGATTTTCAAGTAAATGAATGGGGGATCCCTCAAGCTGTCTGATGTGGAGACA 2700
QY 2899 GACTGAAGCCATTAGACGGGAGCTGCACGGCAGACACAGAGATGTCTAGACATCTTT 2958
DB 2701 GACTGAAGCCATTAGACGGGAGCTGCACGGCAGACACAGAGATGTCTAGACATCTTT 2760
QY 2959 TGGTTGGGACTCCCAATATATTCACCTGAAGTGTTCACGAACAGGATACACACAGT 3018
DB 2761 TGGTTGGGACTCCCAATATATTCACCTGAAGTGTTCACGAACAGGATACACACAGT 2820
QY 3019 TGTGTGATTTGGTGGAGTGTGGTGTATTTTGAATGTTGGTGGGACAACTCCCTT 3078
DB 2821 TGTGTGATTTGGTGGAGTGTGGTGTATTTTGAATGTTGGTGGGACAACTCCCTT 2880
QY 3079 TCTTGGCACAACACCAATATAGAACACAAATGAAGTTATCACTGCGAACAATCTCTTC 3138
DB 2881 TCTTGGCACAACACCAATATAGAACACAAATGAAGTTATCACTGCGAACAATCTCTTC 2940
QY 3139 ACATTCACCAACAGCTTAAGCTCAAGCTTCTGATCTTATTTAACTTTTGGC 3198
DB 2941 ACATTCACCAACAGCTTAAGCTCAAGCTTCTGATCTTATTTAACTTTTGGC 3000
QY 3199 GAGACCCGAAGATCGGTTAGGCAAGAAATGGTGTGATGAATAAAGCTCATCCATTTT 3258
DB 3001 GAGACCCGAAGATCGGTTAGGCAAGAAATGGTGTGATGAATAAAGCTCATCCATTTT 3060
QY 3259 TTAACCAATTTGACTTCCAGTCACTGACAGACAGCTGCTGCTATCATATTCCTAAAA 3318
DB 3061 TTAACCAATTTGACTTCCAGTCACTGACAGACAGCTGCTGCTATCATATTCCTAAAA 3120
QY 3319 TCACACACCCCAACAGATACATCAATTTTGAATGCTGTGATGCTGATAAATATATGGAGTG 3378
DB 3121 TCACACACCCCAACAGATACATCAATTTTGAATGCTGTGATGCTGATAAATATATGGAGTG 3180
QY 3379 ATGATAACGGAAGAAATGTAATGACACTCTCAATGGATGCTATATAAATGGAAGC 3438
DB 3181 ATGATAACGGAAGAAATGTAATGACACTCTCAATGGATGCTATATAAATGGAAGC 3240
QY 3439 ATCTTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTTGAATGACAAATGGCTACC 3498
DB 3241 ATCTTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTTGAATGACAAATGGCTACC 3300
QY 3499 CATATAATATCCGAAGCCTATGTAATGATGATACATTAATTCACAGGCTCAGACAGC 3558
DB 3301 CATATAATATCCGAAGCCTATGTAATGATGATACATTAATTCACAGGCTCAGACAGC 3360
QY 3559 AGTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAATAATCGCGATCTAGTATATG 3618
DB 3361 AGTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAATAATCGCGATCTAGTATATG 3420
QY 3619 TTTAACACACATGATTAATTAATGTAATGAGGATTTGTAAGGGCCTGAAATGGAGGTG 3678
DB 3421 TTTAACACACATGATTAATTAATGTAATGAGGATTTGTAAGGGCCTGAAATGGAGGTG 3480
QY 3679 TTTTGAAGTCTGAGAGTAAATATGCAATATGACAGAGCTATATATGTTGCTCTGT 3738
DB 3481 TTTTGAAGTCTGAGAGTAAATATGCAATATGACAGAGCTATATATGTTGCTCTGT 3540
QY 3739 GTACAAATTTTATTTTCTTAAATATGGAATCTCTTTTAAATGTTTAAATTTTAT -CCA 3797
DB 3541 GTACAAATTTTATTTTCTTAAATATGGAATCTCTTTTAAATGTTTAAATTTTATTTCCCA 3600

QY 3798 GCGTTTAAATCAGTATTTAGAAAAAATTTGTATAAGGAAAGTAAATATGAACTGAAT 3857
DB 3601 GCGTTTAAATCAGTATTTAGAAAAAATTTGTATAAGGAAAGTAAATATGAACTGAAT 3660
QY 3858 ATTATAGTCAGTCTTGTGGTAAAGTACTTAAATAAGTACTTGTGTTTAAAGGA 3917
DB 3661 ATTATAGTCAGTCTTGTGGTAAAGTACTTAAATAAGTACTTGTGTTTAAAGGA 3720
QY 3918 GAACTGTGATCTATTTGTATATATGCTAAATAATTTTAAATAACAGAGTTTGTGAAA 3977
DB 3721 GAACTGTGATCTATTTGTATATATGCTAAATAATTTTAAATAACAGAGTTTGTGAAA 3780
QY 3978 TTTTGT 3984
DB 3781 TTTTGT 3787
RESULT 4
AR201456 AR201456 3213 bp DNA linear PAT 20-APR-2002
LOCUS Sequence 5 from patent US 6359193
DEFINITION
ACCESSION AR201456
VERSION AR201456.1 GI:20252344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3213)
AUTHORS Xu, T., Rao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of late genes
JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
FEATURES Location/Qualifiers
1..3213
source
BASE COUNT 946 a 764 c 696 g 807 t
ORIGIN
Query Match 62.3%; Score 2483; DB 6; Length 3213;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
QY 732 GTCCAGCAATCACTTAACCGCAACAGAGCTGGAAGGTTCTAAAGATCCTTTAGTTCT 791
DB 1 GTCCAGCAATCACTTAACCGCAACAGAGCTGGAAGGTTCTAAAGATCCTTTAGTTCT 60
QY 792 CAGAGGATGCGCGCCACCTAGGAGAAAGTGTGCGCTATCATCTTCTGAGAGTCCCAACTCA 851
DB 61 CAGAGACACGCGCCATCTCTAGGAGAAATGTGTTTATCGTTCTGAAAGCCCAACTCA 120
QY 852 CAGAGATGTAGAGACCTTTGTCTGGATCTGGTATATATCAGCATTTGCTCAAGCTCAC 911
DB 121 CAGGCGGATGTAGAGACCTCTGTCTGGATCGGCAATTCAGCATTTGCTCAAGCTCAC 180
QY 912 CCTAGCAACGACAGAGAGTGAACCCCGCCACACACCTCAAGTAAAGAGGTGTTACTCCT 971
DB 181 CCAAGCAATGGACAGAGTGAACCCCGCCACACCTCAAGTAAAGAGGTGTTACTCCT 240
QY 972 CCACCACTCCAGAGCCAGACTCCCGCTCCAGAGGTACACTCCACCTCCCGCTCA 1031
DB 241 CCACCACTCCAGAGCCAGACTCCCGCTCCAGAGGTACACTCCCGCTCCCGCTCA 300
QY 1032 TGGGAACCAACTCTCAACAAACAGCGTATTCTGGAACATGGAATACGTAATCTCCGA 1091
DB 301 TGGGAACCAACTCTCAACAAACAGCGTATTCTGGAACATGGAATACGTAATCTCCGA 360
QY 1092 ATCTCTCTGCTCCACCTCGGGCATGGCAAGAGGCTATCCTCCACACCTCTCAACACT 1151
DB 361 ATCTCTCTGCTCCACCTCGGGCATGGCAAGAGGCTATCCTCCACACCTCTTACCCT 420
QY 1152 TCCCGCATGATCTCTCTATCAAGGACAGAGGCAATAGTCTCTGTTGGGAGA 1211
DB 421 TCTCCCATGAATCCCGCTTAGCCAGGCTCAGAGGCGCAATAGTCTCTGTTGGGAGA 480

[illegible]

```

FEATURES
source
gene
CDS
    Congress Ave BCM 254D, New Haven, CT 06536, USA
    Location/Qualifiers
        1..3213
        /organism="Mus musculus"
        /db_xref="taxon:10090"
        /tissue_type="brain"
        /dev_stage="newborn"
        <1..3213
        /gene="Lats1"
        <1..2889
        /gene="Lats1"
        /note="putative serine threonine kinase; (MUM) Lats1".
        /codon_start=1
        /product="large tumor suppressor 1"
        /protein_id="AADI6883.1"
        /db_xref="GI:4324436"
        /translation="VQHSINRKQSGWKSGKESLPQRHGPISLGGNVVVRSESPNSQADV
        GRESSGTGIAFAQAPNSGQVRVPPPPVRSVTPPPRGOTTTPPPPPSWME
        PSSQTKRYSGNMEIVISRISFVPQAWQEGYPPPLTSPNPPSQAORALSSVPQGSF
        QITMGSTKFNTPPGVQGGQSDFIHONVPGTSTVTRPPPPYPLTIPVANGQSF
        SALQTGAASAPPFANGVPSQSMVPPNRNSHNMELYININVPGLQTAMPQSSSAPQAQSS
        PSCGHELTQWPIINPVRSNGFNPLGSRASHANSQPSATVTAITPAPIQQPVMKMS
        VLAPLETALATHPSNMQPVQTPPTFSEGTASSVPVPIVAEAPSGVQPPPPYTVY
        KHLHONPSVPRVSVESKCPCKDEQSLPKEDDESKSDGSDGKKEKRLTTSPTIYV
        KNKDEBERESRISQVSPQAFKFMFEQHVNLVSKSHOORLHRKKOLENEMRVLGSLQ
        AQDQRMWLCCKGSNYLTKRAMKDMSEVKITLGIAGFCEVLARKVDTKALYATVK
        TLRRKQVLLRNQVAHVAERDILAEADNENWVRLYYSPQDKNLYFVMDIIPGQDMMS
        LLIRMGFIPENLARFYIAELTAEVSNHVMKGFTHRDIKPNLIDRQGHKLTPDGLC
        TGFRTWHDYSYQSGDHPQSDMSDFSNWQSDPNCRCGRDLRKLPERRAAHOHQRLCA
        SLVGTPNYIAEVLRTGYQLCDWNSGVYILCEMLYQGPFFLAQTPLETKQKVLVIM
        TSLHIPOAKLSPEASDILIKLCRGPEDRLGKNGADEIKAHFPFKTIDFSSDLRQOQS
        SYIPKITHPTDTSNEDVPDPLKWSGDGSEENISDTLSGMYKNGKHPEAHFYEFTFR
        FFDNDGYPVYKPYEYETHSGQSEQDDQDDQTTSDGNNRLVYV"
        764..807
    
```

Query Match	62.3%	Score 2483;	DB 10;	Length 3213;
Best Local Similarity	87.4%;	Pred. No. 0;		
Matches 2769;	Conservative	0;	Mismatches 385;	Indels 15; Gaps
QY	732	GTGCAGCAATCAGTTAAACCGCAACAGAGCTCGAAGGTTCTAAAGAAATCCTTAGTTCCT	791	
DB	1	GTGCAACATTCAAATTAACCGAAACAAAGCTGGAAGGTTCTTAAGAGTCTCTAGTTCCT	60	
QY	792	CAGAGGCATGGCCGCCACTAGAGAAAGTGTGGCTATCATCTTGAGAGTCCCAACTCA	851	
DB	61	CAGACACAGGCCCATCTCTAGAGAAATGTGGTTATCTGTTCTGAAAGCCCAACTCA	120	
QY	852	CAGACAGATGTAGGAAGACCTTTGTCTGATCTGGTATATCAGCATTTCTTCAAGCTCAC	911	
DB	121	CAGGGGATGTAGGAAGACCTCTGTCTGATCGGCAITTCGACATTTGCTCAAGCTCAC	180	
QY	912	CCTAGCAACGGACAGAGAGTGAACCCCCCACCACCACTCAAGTAAGGAGTGTACTTCCT	971	
DB	181	CCAAGCAATGGACAGAGAGTGAACCCCCCACCACCACTCAAGTTAGGAGTGTACTTCCT	240	
QY	972	CCACCACCTCCAAGAGGCGAGACTCCCCCTCCAAGAGGTACAACCTCCACTCCCCCTTCA	1031	
DB	241	CCACCACCTCCGAGAGGCGAGACCCCACTCCCGGAGGCACTCCCTCCGCCCTCA	300	
QY	1032	TGGGAACCAAACTCTCAACAAAGCCGTATTCTGGAACATGGAATACGTAATCTCCCGA	1091	
DB	301	TGGGAACCAAGCTCTCAGACAAAGCCGTACTCTGGGAACATGGAGTAGTACGTAATCTCCCGA	360	
QY	1092	ATCTCTCTGTGCCACTTGGGCGATGCAAGAGGGGTATCCTCCACCACTCTCAACACT	1151	
DB	361	ATCTCTCCCTGTCCACTTGGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCACCT	420	
QY	1152	TCCCCCATGAATCCTCCTTAATCAAGGACAGAGGCAATTAGTCTGTCTGTGGGAGA	1211	
DB	421	TCCTCCATGAATCCCCCTAGCCAGGCTCAGAGGGCCATTAGTCTGTCTGTGGGTAGA	480	
QY	1212	CAACCAATCATGACGAGTCTTAGCAAAATTTAACTTTCCATCAGGAGACCTTGGGAATG	1271	

[illegible]

Db	1558	AAAGAGCTTAACATATATTCGTCTTAAAGGGCTAAAAATGGACAAGTCTATGTTTGTAAAG	1611
Qy	2352	ATAAGACACTAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGATACT	2411
Db	1618	ATAAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTACCAAGAAAAGTCGATACT	1677
Qy	2412	AAGGCTTTGTATGCAACAAAAACTCTTCGAAAGAAAAGATGTTCTTCTTCGAAATCAAGTC	2471
Db	1678	AAAGCTTTGTATGCAACAAAGACTCTTCGAAAGAAAAGACGTTCTGCTCGGAAATCAGGTG	1737
Qy	2472	GCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGCACAATGAATGGGTAGTTGCT	2531
Db	1738	GCTCATGTGAAAGCGGAGGAGATATCCTTAGCAAGAACGCAATGAGTGGTGGTCCCG	1797
Qy	2532	CTATATTATTCATTCCAAGATAAGGACAATTTATACCTTTGTAATGGACTACATCTCCTGGG	2591
Db	1798	CTGTACTACTCTTTCAGGACAAAGCAACTGTACTTTGTGATGGACTACATCTCCTGGG	1857
Qy	2592	GGTGATATGATGAGCCTATTAATTAGAATGGGCATCTTTCAGAAAAGTCTGGCACGATTC	2651
Db	1858	GGGGATATCATGAGCCATTAATTAACAATGGGCATCTTCCCTGAAAAATCTGGCACGATTC	1917
Qy	2652	TACATAGCAGAATACCTTACCTGTGCAGTTGAAAGTGTCAATAAATGGGTTTTATTCATAGA	2711
Db	1918	TACATAGCAGAATTAACCTGTGCAGTTGAAAGTGTCAATAAATGGGTTTTATTCATAGA	1977
Qy	2712	GATATTAACCTGATAAATATTTGATTGATCGGTGATGTTCAVATTAATTAATGACTGACTTT	2771
Db	1978	GATATTAACCTGATAAATTTGATTGACGGTATGGCCATTAATAATGACTGACTTT	2037
Qy	2772	GGCCTCTGCACCTGGCTTCAGATGGACACACGATTCCTAAAGTACTATCAGAGTGGTGACCAT	2831
Db	2038	GGCTCTGCACCTGGCTTCAGATGGACACATGACTCCAAGTACTACCAGAGTGGGATCAC	2097
Qy	2832	CCAGGCAAGATAGCATGGATTTCACTAATGAATGGGGATCCCTCAAGCTGTCGATGT	2891
Db	2098	CCAGGCAAGATAGCATGGATTTCACTAAGATGGGGATCCTTCCCAATTTCTCGGTGT	2157
Qy	2892	GGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCACGACGACGATCTCTAGCA	2951
Db	2158	GGGACAGACTGAAGCCATGAGCGGAGAGCTGCTGCCAGCACGACGATGCTCTAGCC	2217
Qy	2952	CATTCTTGGTGGGACTCCCAATTAATTTGCACCTGAAGTGTGTCTAGCAACAGGATAC	3011
Db	2218	CATTCTCTGGTGGGACTCCCAATTAATTTGCACCTGAAGTGTCTAGCAACAGGATAT	2277
Qy	3012	ACACAGTTGTGTATGGTGGAGTGTGGTGTATTCTTTTGAATATGTTGGTGGGACAA	3071
Db	2278	ACACAGCTGTGTACTGGTGGAGTGTGGTGTATTCTTTTGAATATGTTGGTGGGACAA	2337
Qy	3072	CCTCTCTTCTGGCAACACACCATTTAGAAACACAATGAAGTGTATCACTGGCAACA	3131
Db	2338	CCTCTCTTCTGGCAACAAACCCATTTAGAAACACAATGAAGTGTATCTGGCAACT	2397
Qy	3132	TCTCTTACATTTCCACCACAAGCTAAACCTCAGTCTCTGAAGCTTCTGATCTTATTTAAA	3191
Db	2398	TCTCTACACATCCCTCTCAAGCTGAAGCTGAGTCTCTGAAGCTCTCAGCTCATTTACAAA	2457
Qy	3192	CTTTGGCGAGGACCGAAGATCGCTTAGCAAGAATGGTGTCTGATGAAATAAAGCTCAT	3251
Db	2458	CTGTGTCGAGGACCAAGAACCGCCTCGGCAAGACGGTCTGATGAGATAAAGGCTCAT	2517
Qy	3252	CCATTTTTTAAACAATTCACATTCTCCAGTGCCTTGAGACAGCAGTCTGCTCATACATT	3311
Db	2518	CCATTTTTTAAAGACCATCGATTTCTCTAGTGATCTGAGACAGCAGTCTGCTCATACATC	2577
Qy	3312	CCTAAAAATCACACACCAACAGATACATCAAAATTTTGAATCTCTGTTGATCCTGATAAATTA	3371
Db	2578	CCTAAAAATCAGCATCCACAGATACATCAATTTTCGACCTTGTGATCCTGATAAATTTG	2637
Qy	3372	TGGAGTGATATAACGAGGAGAAAATGTAAATGACACTCTCAATGGATGGTATAAAAAAT	3431
Db	2638	TGGAGCATGGCAGGAGGAGAAAATATCATGTGACACTCTGAGCGGATGGTATAAAAAAT	2697

```

QY 3432 GGAAGCATCTGACATGATCTATGAATTTACCTTCGGAAGGTTTTTTGATGACAAT 3491
Db 2698 GGAAGACCCCGAGCAGCGTTCTTATGATGATTCACCTTCGGAAGGTTTTTTGATGACAAT 2757
QY 3492 GGCTACCATATATATATCCGAAGCCTATTGAATATGAATACATTAATTCACAGGCTCA 3551
Db 2758 GGCTACCATATATATATCCGAAGCCTATTGAATATGAATACATTAATTCACAGGCTCA 2817
QY 3552 GAGCAGCATCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAATAATCCGATCTTA 3611
Db 2818 GAACAACAGTCTGATGAAGATGATCAACACAGGCTCCGATGGAACAAACCGAGATCTTA 2877
QY 3612 GTATATGTTTAAACACACTAGTAATAATGTAATGAGGATTTGTAAGAAGGCGCTGAATG 3671
Db 2878 GTGATGTTTAAATAACATAGGATGATCTGTAATGAGGATTTGTAAGAAGGCGCTGAATG 2933
QY 3672 CGAGGCTGTTTGAAGTCTCGAGAGTAAATATGCAATATGACAGAGCTATATATGTT 3731
Db 2934 CAGGGGTTTTGAAGTTTTGAG---AAATATGCAATGACAGAG---TTTGTG 2985
QY 3732 GCTCTGTGTAATATATTTTCCTAATATATGGAATATGGAATATGGAATATGGAATATGGAAT 3791
Db 2986 GCTCTGTGTAATATATTTTCCTAATATATGGAATATGGAATATGGAATATGGAATATGGAAT 3045
QY 3792 ATCCAGCGCTTTAAATCAGTATTTAGAAAAAATTTGTAAGAAGGCTGAATATGGAAT 3851
Db 3046 ATCCAGCGCTTTAAATCAGTATTTAGAAAAAATTTGTAAGAAGGCTGAATATGGAAT 3105
QY 3852 CTGAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900
Db 3106 CTGATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT

RESULT 6
E24613
LOCUS 2442 bp DNA linear PAT 18-JUN-2001
DEFINITION
warts protein, polynucleotide encoding the protein, antisense
polynucleotide thereof and antibody recognizing the protein.
ACCESSION E24613
VERSION JP 1999089580-A/1.
KEYWORDS
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE
1 (bases 1 to 2442)
Toshihiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
AUTHORS
warts protein, polynucleotide encoding the protein, antisense
TITLE
polynucleotide thereof and antibody recognizing the protein
JOURNAL
SUMITOMO ELECTRIC IND LTD
COMMENT
OS Unidentified
PN JP 1999089580-A/1
PD 06-APR-1999
PR 24-SEP-1997 JP 1997258689
PI TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
YASUYUKI NISHIYAMA,
PI HIDEYUKI SAVA
PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
G01N33/532, C12P21/02,
PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
(C12N15/00, PC C12R1:91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT Source 1..2442
FT Location/Qualifiers
1..2442
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 777 a 528 c 491 g 646 t

```

```

ORIGIN
Query Match 60.8%; Score 2422; DB 6; Length 2442;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2436; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1062 TCTGGAACATGGAATACGTAATCTCCCGAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 1121
Db 1 TCTGGAACATGGAATACGTAATCTCCCGAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 60
QY 1122 GAGGGCTATCTCCACACCTCTCAACACTTCCCGCATGAATCTCCCTCAATCAAGCAGAG 1181
Db 61 GAGGGCTATCTCCACACCTTTTCAACACTTCCCGCATGAATCTCCCTCAATCAAGCAGAG 120
QY 1182 AGAGGCTATGTTCTGTTGTCGAGACCAACCAATCATCATGCGAGAGTCTTACGAAA 1241
Db 121 AGAGGCTATGTTCTGTTGTCGAGACCAACCAATCATCATGCGAGAGTCTTACGAAA 180
QY 1242 TTTAACTTTCCATCAGGAGACCTGGAATGCAAGTGGTACTGGACAAACTGATTTCTATG 1301
Db 181 TTTAACTTTCCATCAGGAGACCTGGAATGCAAGTGGTACTGGACAAACTGATTTCTATG 240
QY 1302 ATACACCAAAATCTTCTGCTGCTGCACTGTGAATCGGAGCCACCACTCCATATCT 1361
Db 241 ATACACCAAAATCTTCTGCTGCTGCACTGTGAATCGGAGCCACCACTCCATATCT 300
QY 1362 CTGACAGCAGCTAATGGACAAAGCCCTTCTGTTTACAAACAGGGGATCTGCTGCTCT 1421
Db 301 CTGACAGCAGCTAATGGACAAAGCCCTTCTGTTTACAAACAGGGGATCTGCTGCTCT 360
QY 1422 TCGTCATATCAAAATGGAAGTATTTCTCAGTCTATGATGTCGCAACAGAAATAGTCAT 1481
Db 361 TCGTCATATCAAAATGGAAGTATTTCTCAGTCTATGATGTCGCAACAGAAATAGTCAT 420
QY 1482 AACATGGAACATATATAACATTAGTGTACCTGGAGCTGCAAAACAAATTTGGCTCAGTCAT 1541
Db 421 AACATGGAACATATATAACATTAGTGTACCTGGAGCTGCAAAACAAATTTGGCTCAGTCAT 480
QY 1542 TCTGCTCCAGCCAGTCATCCCGAGCAGTGGGATGAATCTCTACATGGCAACCTAAC 1601
Db 481 TCTGCTCCAGCCAGTCATCCCGAGCAGTGGGATGAATCTCTACATGGCAACCTAAC 540
QY 1602 ATACCACTGAGGTCAAAATTTCTTTAATAACCCATTAGAAATAGAGCAAGTCACTCTGCT 1661
Db 541 ATACCACTGAGGTCAAAATTTCTTTAATAACCCATTAGAAATAGAGCAAGTCACTCTGCT 600
QY 1662 AATTCTCAGCCTTCTGCTACACAGTCACTGCAATTTACACAGCTCTCTTATCAACAGCT 1721
Db 601 AATTCTCAGCCTTCTGCTACACAGTCACTGCAATTTACACAGCTCTCTTATCAACAGCT 660
QY 1722 GTGAAAAGTATGCTGTTAATAACAGGCTACAGACTGCTTTAGCACCTACACACCT 1781
Db 661 GTGAAAAGTATGCTGTTAATAACAGGCTACAGACTGCTTTAGCACCTACACACCT 720
QY 1782 TCTTGGATACACAGCAATTTCAAACTGTTCAACCCAGTCTCTTTCTGAGGGAACCGCT 1841
Db 721 TCTTGGATACACAGCAATTTCAAACTGTTCAACCCAGTCTCTTTCTGAGGGAACCGCT 780
QY 1842 TCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAACTCAAGCAAGCACCACCA 1901
Db 781 TCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAACTCAAGCAAGCACCACCA 840
QY 1902 CCCTACCAAAACATCTGCTGACCAAAACCCATCTCTCTCCATAGGAGTCAATCAGT 1961
Db 841 CCCTACCAAAACATCTGCTGACCAAAACCCATCTCTCTCCATAGGAGTCAATCAGT 900
QY 1962 AAGCCTAGCAAGAGGATCAGCCAGCTTGCCCAAGGAAGATGAGTGAAGAGGTTAT 2021
Db 901 AAGCCTAGCAAGAGGATCAGCCAGCTTGCCCAAGGAAGATGAGTGAAGAGGTTAT 960
QY 2022 GAAATGTTGATGTTGGGATTAAGAAAGAAACAGATTACAACTTACCTTACTTACTGTT 2081
Db 961 GAAATGTTGATGTTGGGATTAAGAAAGAAACAGATTACAACTTACCTTACTTACTGTT 1020

```

QY 2082 AGGAAACAAAGAGATGAAGAGCGAAGGAATCTCGTATTCAAAGTTATTCTCTCAA 2141
DB 1021 AGGAAACAAAGAGATGAAGAGCGAAGGAATCTCGTATTCAAAGTTATTCTCTCAA 1080
QY 2142 GCATTAAATCTTTATGGAGCAACATGTAGAAAATGACTCAAAATCTCATCAGCAGGT 2201
DB 1081 GCATTAAATCTTTATGGAGCAACATGTAGAAAATGACTCAAAATCTCATCAGCAGGT 1140
QY 2202 CTACATCTTAAACAAATAGAGAAATGAAGATGATGCGGGTGGATATTCTCAAGATGCC 2261
DB 1141 CTACATCTTAAACAAATAGAGAAATGAAGATGATGCGGGTGGATATTCTCAAGATGCC 1200
QY 2262 CAGGATCAATGAGAAAGATGCTTTGCCAAAAGAAATCTAATTACATCCGCTCTTAAAGG 2321
DB 1201 CAGGATCAATGAGAAAGATGCTTTGCCAAAAGAAATCTAATTACATCCGCTCTTAAAGG 1260
QY 2322 GCTAAATGGCAAGTCTATGTTGTGGAAGATAAGACACACATAGGAGCAATTTGGT 2381
DB 1261 GCTAAATGGCAAGTCTATGTTGTGGAAGATAAGACACACATAGGAGCAATTTGGT 1320
QY 2382 GAAGTCTCTAGCAAGAAAGTACTACTAAGGCTTTGTATGCAACAAACTCTTCCA 2441
DB 1321 GAAGTCTCTAGCAAGAAAGTACTACTAAGGCTTTGTATGCAACAAACTCTTCCA 1380
QY 2442 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTG 2501
DB 1381 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTG 1440
QY 2502 GCTGAAGTGCACATGAATGAATGGTGTGCTGTATATTTATTCATTCACAAAGATAGGACAA 2561
DB 1441 GCTGAAGTGCACATGAATGAATGGTGTGCTGTATATTTATTCATTCACAAAGATAGGACAA 1500
QY 2562 TTATACCTTGAATGGACTACATTCCTGGGGTGATATGATGAGCCTATTATTAAGATG 2621
DB 1501 TTATACCTTGAATGGACTACATTCCTGGGGTGATATGATGAGCCTATTATTAAGATG 1560
QY 2622 GGCATCTTTCCAGAAAGTCTGACGATTTCTACATGAGAGAACTTACCTGTGCAGTTGAA 2681
DB 1561 GGCATCTTTCCAGAAAGTCTGACGATTTCTACATGAGAGAACTTACCTGTGCAGTTGAA 1620
QY 2682 AGTGTTCATAAATGGGTTTTATTTCATGAGATATTAACCTGATATTTTGAATGAT 2741
DB 1621 AGTGTTCATAAATGGGTTTTATTTCATGAGATATTAACCTGATATTTTGAATGAT 1680
QY 2742 CGTATGTCATATTAATTAATGACTGACCTTGCCTCTGCCTGCTGAGATGGACAC 2801
DB 1681 CGTATGTCATATTAATTAATGACTGACCTTGCCTCTGCCTGCTGAGATGGACAC 1740
QY 2802 GATTTAAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAAGTAA 2861
DB 1741 GATTTAAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAAGTAA 1800
QY 2862 GAATGGGGGATCCCTCAAGCTGTCGATGTGAGACAGACTGAAGCCATTAGAGCGGAGA 2921
DB 1801 GAATGGGGGATCCCTCAAGCTGTCGATGTGAGACAGACTGAAGCCATTAGAGCGGAGA 1860
QY 2922 GCTGACCCGACGACCGGATGCTAGCACATCTTTGGTGGACCTCCCAATTTATTT 2981
DB 1861 GCTGACCCGACGACCGGATGCTAGCACATCTTTGGTGGACCTCCCAATTTATTT 1920
QY 2982 GCACCTGAAGTGTGTTACGAAACAGGATACACAGTTGTGTGATTTGGTGGAGTCTTGGT 3041
DB 1921 GCACCTGAAGTGTGTTACGAAACAGGATACACAGTTGTGTGATTTGGTGGAGTCTTGGT 1980
QY 3042 GTTATCTTTTGAATGTTGGTGGACCACTCTCTTTTGGCAACAAACCAATTAGAA 3101
DB 1981 GTTATCTTTTGAATGTTGGTGGACCACTCTCTTTTGGCAACAAACCAATTAGAA 2040
QY 3102 ACACAAATGAAGTTATCAACTGGCAACATCTCTTCATCTCCACCAAGCTTAACTC 3161
DB 2041 ACACAAATGAAGTTATCAACTGGCAACATCTCTTCATCTCCACCAAGCTTAACTC 2100

QY 3162 AGTCTCTGAAGCTTCTGATCTTATTATAAATTTGCGAGGACCCGAAGATCGCTTAGGC 3221
DB 2101 AGTCTCTGAAGCTTCTGATCTTATTATAAATTTGCGAGGACCCGAAGATCGCTTAGGC 2160
QY 3222 AAGAATGGTGTGATGAATAAAGCTCATCTTTTAAAAACAATTTGCTTCCAGT 3281
DB 2161 AAGAATGGTGTGATGAATAAAGCTCATCTTTTAAAAACAATTTGCTTCCAGT 2220
QY 3282 GACCTGAGACAGCACTGCTTCTATACATCTCTTAAATACACACACCCACACATACATCA 3341
DB 2221 GACCTGAGACAGCACTGCTTCTATACATCTCTTAAATACACACACCCACACATACATCA 2280
QY 3342 AATTTTGTATCTGTTGATCTCTGATAAATATGAGTGTATGATAAGGAGAAATGTA 3401
DB 2281 AATTTTGTATCTGTTGATCTCTGATAAATATGAGTGTATGATAAGGAGAAATGTA 2340
QY 3402 AATGACATCTCAAT- GGATGGTATAAATAAGGAGGATCTGACATGCTTCTATGA 3460
DB 2341 AATGACATCTCAATGGGATGATAAATAAGGAGGATCTGACATGCTTCTATGA 2400
QY 3461 ATTTACCTCCGAAGGTTTTTTTGATGACAAATGGCTACCATA 3502
DB 2401 ATTTACCTCCGAAGGTTTTTTTGATGACAAATGGCTACCATA 2442

RESULT 7

AL583963/C

LOCUS AL583963 Human DNA sequence from clone RPI-203A15 on chromosome 6, complete sequence.
DEFINITION

ACCESSION AL583963

VERSION AL583963.11 GI:14547269

KEYWORDS HTG.

SOURCE human.

ORGANISM

human.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 80146)

Hammond, S.

Direct Submission

Submitted (23-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,

CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

requests: clonerequest@sanger.ac.uk

On Jun 25, 2001 this sequence version replaced gi:14133045.

During sequence assembly data is compared from overlapping clones.

Where differences are found these are annotated as variations

together with a note of the overlapping clone name. Note that the

variation annotation may not be found in the sequence submission

corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all

regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest. The following

abbreviations are used to associate primary accession numbers given

in the feature table with their source databases: Em, EMBL; Sw,

SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP

database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence

was generated from part of bacterial clone contigs of human

chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping

group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr6

RPI-203A15 is from the library RPI-1 constructed by the group of

Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm

VECTOR: pcYFAC2

IMPORTANT: This sequence is not the entire insert of clone

RPI-203A15 it may be shorter because we sequence overlapping

sections only once, except for a 100 base overlap.

The true left end of clone RP1-317N9 is at 78147 in this sequence.
The true right end of clone RP11-703H16 is at 2000 in this sequence.

FEATURES

Location/Qualifiers

1. 80146
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP1-203A15"
/clone_lib="RPCL1"

repeat_region

/note="L1M5 repeat: matches 7790. 7931 of consensus"

repeat_region

/note="L2 repeat: matches 2234. 2359 of consensus"

misc_feature

2135. 2886
/note="CpG island"

repeat_region

/evidence="not_experimental"

repeat_region

2728. 3050
/note="AluJb repeat: matches 1. 303 of consensus"

repeat_region

3174. 3476
/note="AluJ repeat: matches 2. 303 of consensus"

repeat_region

3512. 3956
/note="L2 repeat: matches 2264. 2750 of consensus"

repeat_region

3975. 4127
/note="AluSg/x repeat: matches 145. 297 of consensus"

repeat_region

4152. 4469
/note="AluSx repeat: matches 1. 304 of consensus"

repeat_region

4470. 4856
/note="HURP3 repeat: matches 6797. 7185 of consensus"

repeat_region

4850. 6348
/note="HURP3 repeat: matches 5120. 6625 of consensus"

repeat_region

6349. 6647
/note="AluSg repeat: matches 1. 301 of consensus"

repeat_region

6648. 6815
/note="HURP3 repeat: matches 4954. 5120 of consensus"

repeat_region

6840. 6911
/note="HURP3 repeat: matches 4691. 4762 of consensus"

repeat_region

6912. 6953
/note="L1P repeat: matches 1790. 1825 of consensus"

repeat_region

6974. 7347
/note="L1P repeat: matches 742. 701 of consensus"

repeat_region

7349. 7384
/note="HURP3 repeat: matches 4568. 4942 of consensus"

repeat_region

7382. 7745
/note="L1P repeat: matches 1590. 1958 of consensus"

repeat_region

7745. 8418
/note="L1M2 repeat: matches -84. 638 of consensus"

repeat_region

8458. 8921
/note="L1M2 repeat: matches -677. -198 of consensus"

repeat_region

9470. 9751
/note="AluSx repeat: matches 12. 293 of consensus"

repeat_region

9941. 10043
/note="AluJb repeat: matches 210. 306 of consensus"

repeat_region

10059. 10260
/note="AluSg/x repeat: matches 108. 309 of consensus"

repeat_region

10295. 10994
/note="L1P1 repeat: matches 4004. 4703 of consensus"

repeat_region

10995. 11289
/note="AluSg repeat: matches 1. 301 of consensus"

repeat_region

11290. 12343
/note="L1P1 repeat: matches 4703. 5780 of consensus"

repeat_region

12344. 12479
/note="AluSx repeat: matches 1. 142 of consensus"

repeat_region

12480. 12769
/note="AluJ repeat: matches 1. 290 of consensus"

repeat_region

12770. 12957
/note="AluSg repeat: matches 142. 309 of consensus"

repeat_region

12958. 13331
/note="L1P1 repeat: matches 5779. 6152 of consensus"

repeat_region

13382. 13962
/note="L1M4A repeat: matches 3027. 3603 of consensus"

repeat_region

13963. 14255
/note="AluSc repeat: matches 1. 306 of consensus"

14256. 15963
/note="L1M4A repeat: matches 3603. 5331 of consensus"

15966. 16070
/note="FIAM_C repeat: matches 1. 105 of consensus"

16076. 16112
/note="Alu repeat: matches 79. 160 of consensus"

16122. 17029
/note="L1M4A repeat: matches 5319. 6300 of consensus"

17080. 17261
/note="MER74A repeat: matches 1. 171 of consensus"

17490. 17788
/note="AluSg repeat: matches 1. 296 of consensus"

17966. 18263
/note="AluJ repeat: matches 1. 298 of consensus"

18502. 18551
/note="25 copies 2 mer ta 100% conserved"

18552. 18818
/note="AluSg repeat: matches 34. 303 of consensus"

18929. 19117
/note="MIR repeat: matches 41. 241 of consensus"

19252. 19557
/note="AluSg repeat: matches 1. 306 of consensus"

19558. 19843
/note="AluJ repeat: matches 12. 294 of consensus"

20971. 21270
/note="AluSx repeat: matches 1. 300 of consensus"

21302. 21391
/note="FRAM/FAM repeat: matches 4. 72 of consensus"

22015. 22173
/note="MER5B repeat: matches 7. 176 of consensus"

22419. 23403
/note="CpG island"

22531. 22618
/evidence="not_experimental"

23378. 23453
/note="44 copies 2 mer cc 64% conserved"

23686. 23896
/note="MER5A repeat: matches 108. 185 of consensus"

24365. 24495
/note="FIAM_C repeat: matches 2. 125 of consensus"

24659. 24859
/note="MER3 repeat: matches 2. 209 of consensus"

25072. 25249
/note="MER3 repeat: matches 118. 290 of consensus"

25256. 25572
/note="AluSg repeat: matches 1. 309 of consensus"

25575. 25874
/note="AluSx repeat: matches 1. 299 of consensus"

26374. 26417
/note="22 copies 2 mer ta 100% conserved"

27253. 27549
/note="AluSg repeat: matches 1. 298 of consensus"

27551. 27631
/note="LIME repeat: matches 5699. 5778 of consensus"

27852. 27915
/note="L2 repeat: matches 2203. 2267 of consensus"

28020. 28320
/note="AluSg repeat: matches 1. 305 of consensus"

28891. 29184
/note="AluSg repeat: matches 1. 292 of consensus"

29260. 29548
/note="AluSg repeat: matches 1. 305 of consensus"

29567. 29652
/note="AluJ/FAM repeat: matches 202. 287 of consensus"

29818. 30056
/note="AluSg repeat: matches 1. 293 of consensus"

30267. 30570
/note="AluSg repeat: matches 1. 302 of consensus"

30581. 30762
/note="LIME repeat: matches 2353. 2196 of consensus"

30827. 31119
/note="AluJb repeat: matches 1. 312 of consensus"

repeat_region	31127..31259	/note="LIMEC repeat: matches 2112..2246 of consensus"
repeat_region	31453..31645	/note="LIMEC repeat: matches 1407..1614 of consensus"
repeat_region	32228..32673	/note="LIMEC repeat: matches 100..541 of consensus"
repeat_region	32688..32820	/note="LIMEC repeat: matches 1..132 of consensus"
Query Match	38.1%;	Score 1518.2; DB 9; Length 80146;
Best Local Similarity	99.5%;	Pred. No. 0;
Matches 1523;	Conservative 0;	Mismatches 8; Indels 0; Gaps 0;
QY 725	AGGGAATGTGACCAATCAGTTAAACCGCAACAGAGCTGGAAAGGTTCTAAAGAACCTT	784
DB 57333	AGGGAATGTGACCAATCAGTTAAACCGCAACAGAGCTGGAAAGGTTCTAAAGAACCTT	57274
QY 785	AGTTCTCAGAGGATGGGCGCCACTAGGAGAAAGTGTGGCCCTATCTTCAGAGTCC	844
DB 57273	AGTTCTCAGAGGATGGGCGCCACTAGGAGAAAGTGTGGCCCTATCTTCAGAGTCC	57214
QY 845	CAACTCACAGACAGATGTAGGAGACCTTGTCTGCATCTGGTATATACGATTTGTTCA	904
DB 57213	CAACTCACAGACAGATGTAGGAGACCTTGTCTGCATCTGGTATATACGATTTGTTCA	57154
QY 905	AGCTCACCTTAGCAACGGACAGAGAGTGAACCCGCCACACCACTCAAGTAAGGAGTGT	964
DB 57153	AGCTCACCTTAGCAACGGACAGAGAGTGAACCCGCCACACCACTCAAGTAAGGAGTGT	57094
QY 965	TACTCTCCACCACTTCCCAAGAGGCGAGACTCCGCCCTCCCAAGAGGTACAACTCCACCTCC	1024
DB 57093	TACTCTCCACCACTTCCCAAGAGGCGAGACTCCGCCCTCCCAAGAGGTACAACTCCACCTCC	57034
QY 1025	CCCTTACCTGGACCAACCACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAAT	1084
DB 57033	CCCTTACCTGGACCAACCACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAAT	56974
QY 1085	CTCCCAATCTCTCTCTGCTCCCACTGGGCGATGGCAAGGGGTATCTCTCCACCACTCT	1144
DB 56973	CTCCCAATCTCTCTCTGCTCCCACTGGGCGATGGCAAGGGGTATCTCTCCACCACTCT	56914
QY 1145	CAACACTTCCCACTGAATCTCTCTATCAAGGACAGAGGAGGATAGTTCTCTCTCTGT	1204
DB 56913	CAACACTTCCCACTGAATCTCTCTATCAAGGACAGAGGAGGATAGTTCTCTCTCTGT	56854
QY 1205	TGSCACAGCAACCAATCATCATGACAGAGTTCTAGCAAAATTAACCTTCCATCAGGAGACC	1264
DB 56853	TGSCACAGCAACCAATCATCATGACAGAGTTCTAGCAAAATTAACCTTCCATCAGGAGACC	56794
QY 1265	TGGAATGCAAGATGGTACTGGCAAACTGATTTTCATGATACACCAAAATGTTCTCCCTGC	1324
DB 56793	TGGAATGCAAGATGGTACTGGCAAACTGATTTTCATGATACACCAAAATGTTCTCCCTGC	56734
QY 1325	TGGCACTGTGAATGCGGAGCCACCACTCCATATCTCTGACAGAGGATTAATGGACAAAG	1384
DB 56733	TGGCACTGTGAATGCGGAGCCACCACTCCATATCTCTGACAGAGGATTAATGGACAAAG	56674
QY 1385	CCCTTCTGCTTTTCAACAGGGGATCTGCTGCTCTCTGCTATATACAAATGGAAGTAT	1444
DB 56673	CCCTTCTGCTTTTCAACAGGGGATCTGCTGCTCTCTGCTATATACAAATGGAAGTAT	56614
QY 1445	TCTCAGTCTATGATGGTCCAAACAGAAATAGTCTATACATGGAATATATACAAATAG	1504
DB 56613	TCTCAGTCTATGATGGTCCAAACAGAAATAGTCTATACATGGAATATATACAAATAG	56554
QY 1505	TGTACTGAGCTGCAAAACAAATTTGGCCCTCAGTCACTCTCTGCTCCAGCCAGTCATCCCC	1564
DB 56553	TGTACTGAGCTGCAAAACAAATTTGGCCCTCAGTCACTCTCTGCTCCAGCCAGTCATCCCC	56494
QY 1565	GAGCACTGGGCGATGAATCCCTACATGGCAACCTACATACCACTGAGTCAAAATCTTTT	1624
DB 56493	GAGCACTGGGCGATGAATCCCTACATGGCAACCTACATACCACTGAGTCAAAATCTTTT	56434

QY 1625	TAATAACCCATTAGGAATAGAGCAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCTACAAC	1684
DB 56433	TAATAACCCATTAGGAATAGAGCAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCTACAAC	56374
QY 1685	AGTCACTGCAATTTACACCACTCTTATTAACAGCCCTGTGAAAGATGCGTGTATTTAA	1744
DB 56373	AGTCACTGCAATTTACACCACTCTTATTAACAGCCCTGTGAAAGATGCGTGTATTTAA	56314
QY 1745	ACCAGAGCTACAGACTGCTTTAGCACTACACACCCCTTCTTGATACCAAGCATTTCA	1804
DB 56313	ACCAGAGCTACAGACTGCTTTAGCACTACACACCCCTTCTTGATACCAAGCATTTCA	56254
QY 1805	AACCTGTTCAACCCAGCTCTTTCTCTGAGGAACCGCTTCAATGTGACTGTGATGCCACC	1864
DB 56253	AACCTGTTCAACCCAGCTCTTTCTCTGAGGAACCGCTTCAATGTGACTGTGATGCCACC	56194
QY 1865	TGTTGCTGAAGCTTCAAACTATCAAGGACCAACCACTTCAACCACTTCAACCACTTCA	1924
DB 56193	TGTTGCTGAAGCTTCAAACTATCAAGGACCAACCACTTCAACCACTTCAACCACTTCA	56134
QY 1925	CAAAACCCATCTGTTCTCTCATACAGTCAATCAGTAAAGCTTCAACCACTTCAACCACTT	1984
DB 56133	CAAAACCCATCTGTTCTCTCATACAGTCAATCAGTAAAGCTTCAACCACTTCAACCACTT	56074
QY 1985	AGCTTTCGCCAAGGAAGAGTGAAGAGTGTATGAAATGTTGATAGTGGGATAA	2044
DB 56073	AGCTTTCGCCAAGGAAGAGTGAAGAGTGTATGAAATGTTGATAGTGGGATAA	56014
QY 2045	AGAAAGAAACAGATTACAACTTCACTTATCTGTTAGGAAACCAAGAAAGATGAAGA	2104
DB 56013	AGAAAGAAACAGATTACAACTTCACTTATCTGTTAGGAAACCAAGAAAGATGAAGA	55954
QY 2105	GCAAGGGATCTGCTTATCAAGTGTATCTCTCCTCAGCAATTTAAATTTCTTTTGGGCA	2164
DB 55953	GCAAGGGATCTGCTTATCAAGTGTATCTCTCCTCAGCAATTTAAATTTCTTTTGGGCA	55894
QY 2165	ACATGTAGAAATGTACTCAATCTCATCAGCAGCGTCTACATCGTAAACCAATTAGA	2224
DB 55893	ACATGTAGAAATGTACTCAATCTCATCAGCAGCGTCTACATCGTAAACCAATTAGA	55834
QY 2225	GAATGAATGATCGGGTGGATTATCTCAA	2255
DB 55833	GAATGAATGATCGGGTGGATTATCTCAA	55803

RESULT 8	E24614	1374 bp	DNA	linear	PAT 18-JUN-2001
LOCUS	E24614				
DEFINITION	warts Protein, polynucleotide encoding the protein, antisense polynucleotide thereof and antibody recognizing the protein.				
ACCESSION	E24614				
VERSION	E24614.1	GI:13024647			
KEYWORDS	JP 1999089580-A/2.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1374)				
AUTHORS	Toshihiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.				
TITLE	warts Protein, polynucleotide encoding the protein, antisense polynucleotide thereof and antibody recognizing the protein				
JOURNAL	Patent: JP 1999089580-A 2 06-APR-1999;				
COMMENT	SUMITOMO ELECTRIC IND LTD				
	OS	Unidentified			
	PN	JP 1999089580-A/2			
	PD	06-APR-1999			
	PF	24-SEP-1997 JP 1997258689			
	PI	TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI YASUYUKI NISHIYAMA, PI HIDEYUKI SAYA PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC G01N33/532/C12P21/02, PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,			

(C12N15/00, PC C12R1:91)	
CC Strandedness: Double;	
FH Key	
FT source	
FEATURES	
source	Location/Qualifiers
	1..1374
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	442 a 252 c 288 g 392 t
ORIGIN	
Query Match	
Best Local Similarity 34.18; Score 1360.4; DB 6; Length 1374;	
Matches 1372: Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
Qy	2130 TATTCTCTCAAGCATTTAAATCTTTATGAGCAACATGTAGAAATGTACTCAAAATCT 2189
Db	1 TATTCTCTCAAGCATTTAAATCTTTATGAGCAACATGTAGAAATGTACTCAAAATCT 60
Qy	2190 CATCAGAGCGCTACATCGTAAACAAATAGAAATGAATGATCGGGTTGGATTA 2249
Db	61 CATCAGAGCGCTACATCGTAAACAAATAGAAATGAATGATCGGGTTGGATTA 120
Qy	2250 TCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAAATCTAAATACATC 2309
Db	121 TCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAAATCTAAATACATC 180
Qy	2310 CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATGAAGACACTAGGAATA 2369
Db	181 CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATGAAGACACTAGGAATA 240
Qy	2370 GGAGCATTTGGTGAAGTCTGTACGAGAAAGTACTAGGCTTTGTATGCAACA 2429
Db	241 GGAGCATTTGGTGAAGTCTGTACGAGAAAGTACTAGGCTTTGTATGCAACA 300
Qy	2430 AAAACTCTTGAAGAAAGATGTTCTTCTGAAATCAAGTCGCTCATGTTAAGGCTGAG 2489
Db	301 AAAACTCTTGAAGAAAGATGTTCTTCTGAAATCAAGTCGCTCATGTTAAGGCTGAG 360
Qy	2490 AGAGATATCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGCTATATATTCATTCAC 2549
Db	361 AGAGATATCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGCTATATATTCATTCAC 420
Qy	2550 GATAAGGCAATTTATCTTTGTAATGACTACATTCCTGGGGGTGATATGATGAGCCTA 2609
Db	421 GATAAGGCAATTTATCTTTGTAATGACTACATTCCTGGGGGTGATATGATGAGCCTA 480
Qy	2610 TTAATTAGAATGGGCATCTTTCCGAAAGTCTGCGACGATTCATAGCAGAACTTACC 2669
Db	481 TTAATTAGAATGGGCATCTTTCCGAAAGTCTGCGACGATTCATAGCAGAACTTACC 540
Qy	2670 TGTGACATTTGAAGTTCATATAAATGGTTTATTCATAGAGATTTAAACCTGATAT 2729
Db	541 TGTGACATTTGAAGTTCATATAAATGGTTTATTCATAGAGATTTAAACCTGATAT 600
Qy	2730 ATTTCATGATGCTGATGGTTCATATTAATTAAGTACTGACTTTGGCCTCTGCACTGGCTC 2789
Db	601 ATTTCATGATGCTGATGGTTCATATTAATTAAGTACTGACTTTGGCCTCTGCACTGGCTC 660
Qy	2790 AGATGACACAGCTTCTAAGTACTATCAGAGTGTGACCATCCACGGCAAGATAGCATG 2849
Db	661 AGATGACACAGCTTCTAAGTACTATCAGAGTGTGACCATCCACGGCAAGATAGCATG 720
Qy	2850 GATTTCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCA 2909
Db	721 GATTTCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCA 780
Qy	2910 TTAGACGGGAGAGCTGCACGCCAGCAGCAGCATGCTTAGCACATCTTTGGTGGGACT 2969
Db	781 TTAGACGGGAGAGCTGCACGCCAGCAGCAGCATGCTTAGCACATCTTTGGTGGGACT 840

RESULT 9	
AF119846/c	
ID	AF119846 standard; RNA; HUM; 1807 BP.
XX	AF119846;
AC	AF119846;
XX	AF119846.1
SV	AF119846.1
XX	12-MAY-2000 (Rel. 63, Created)
DT	09-MAY-2001 (Rel. 67, Last updated, Version 2)
XX	Homo sapiens PRO1474 mRNA, complete cds.
DE	
XX	
KW	
OS	Homo sapiens (human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX	[1]
RP	1-1807
RA	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RT	"Functional prediction of the coding sequences of 79 new genes deduced by
RT	analysis of cDNA clones from human fetal liver";
RL	Unpublished.
XX	[2]
RP	1-1807
RA	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RT	Submitted (13-JAN-1999) to the EMBL/GenBank/DBJ databases.
RL	Department of Experimental Hematology, Institute of Radiation Medicine,
RL	Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
XX	SPTREMBL; Q9P199; Q9P199.
DR	
XX	

QY	723	CCAGGGAATGTGCAGCAATCAGTTTAACCGGAAAACAGAGCTGCGAAAGCTTCTTAAGAATCC	782
Db	214	CCAGGGAATGTGCAGCAATCAGTTTAACCGGAAAACAGAGCTGCGAAAGCTTCTTAAGAATCC	155
QY	783	TTAGTTCTCTCAGAGGATGGCCGCCCACTAGGAGAAAGTGTGGCTCATCATCTTGAGAGT	842
Db	154	TTAGTTCTCTCAGAGGATGGCCGCCCACTAGGAGAAAGTGTGGCTCATCATCTTGAGAGT	95
QY	843	CCCAACTCACAGACAGATGTAGGAAGACCTTGTCTCGATCTCGTATATCAGCATTTGTT	902
Db	94	CCCAACTCACAGACAGATGTAGGAAGACCTTGTCTCGATCTCGTATATCAGCATTTGTT	35
QY	903	CAAGCTCACCTAGCAACGGACAGAGAGTGAACC	936
Db	34	CAAGCTCACCTAGCAACGGACAGAGAGTGAACC	1

RESULT 10
 AC102712/c
 LOCUS Mus musculus clone RP24-258P4, WORKING DRAFT SEQUENCE, 17 unordered pieces.
 DEFINITION AC102712 164490 bp DNA linear HTG 21-AUG-2002
 AC102712.2 GI:22381704
 HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 KEYWORDS house mouse.
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 164490)
 Birren,B., Nusbaum,C. and Lander,E.
 Title
 Mus musculus, clone RP24-258P4
 Unpublished
 2 (bases 1 to 164490)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J., Chazaro,B., Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Glende,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Jones,C., Kearfott,A., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamarez,R., Landers,T., Lechoczky,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Majoor,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164490)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliiev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,C., MacLean,C., Macdonald,P., Majoor,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17060822.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L19333

Center clone name: 258_P4

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 158929 bases at least Q40

Consensus quality: 161241 bases at least Q30

Consensus quality: 162174 bases at least Q20

Insert size: 176000; agarose-fp

Quality coverage: 7.2 in Q20 bases; agarose-fp

Quality coverage: 7.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1303: contig of 1303 bp in length
1304 1403: gap of 100 bp
1404 2163: contig of 760 bp in length
2164 2263: gap of 100 bp
2264 3531: contig of 1268 bp in length
3532 3631: gap of 100 bp
3632 4716: contig of 1085 bp in length
4717 4816: gap of 100 bp
4817 8090: contig of 3274 bp in length
8091 8190: gap of 100 bp
8191 11116: contig of 2926 bp in length
11117 11216: gap of 100 bp
11217 41036: contig of 29820 bp in length
41037 41136: gap of 100 bp
41137 45062: contig of 3926 bp in length
45063 45162: gap of 100 bp
45163 48629: contig of 3467 bp in length
48630 48729: gap of 100 bp
48730 57775: contig of 9046 bp in length
57776 57875: gap of 100 bp
57876 70851: contig of 12976 bp in length
70852 70951: gap of 100 bp
70952 85997: contig of 15046 bp in length
85998 86097: gap of 100 bp
86098 104897: contig of 18800 bp in length
104898 104997: gap of 100 bp
104998 120197: contig of 15200 bp in length
120198 120297: gap of 100 bp
120298 136805: contig of 16508 bp in length
136806 136905: gap of 100 bp
136906 160104: contig of 23199 bp in length
160105 160204: gap of 100 bp

* 160205 164490: contig of 4286 bp in length.

FEATURES

Source

Location/Qualifiers
1. 164490
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-258P4"
/clone_lib="RPCI-24 Male Mouse BAC"

misc_feature

1. 1303

/note="assembly_fragment"

1404. 2163

/note="assembly_fragment"

2264. 3531

/note="assembly_fragment"

3632. 4716

/note="assembly_fragment"

4817. 8090

/note="assembly_fragment"

8191. 11116

/note="assembly_fragment"

11217. 41036

/note="assembly_fragment"

41137. 45062

/note="assembly_fragment"

45163. 48629

/note="assembly_fragment"

48730. 57775

/note="assembly_fragment"

57876. 70851

/note="assembly_fragment"

70952. 85997

/note="assembly_fragment"

86098. 104897

/note="assembly_fragment"

104998. 120197

/note="assembly_fragment"

120298. 136805

/note="assembly_fragment"

136906. 160104

/note="assembly_fragment"

160205. 164490

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 47330 a 35286 c 34742 g 45505 t 1627 others

ORIGIN

Query Match 21.3%; Score 848.4; DB 2; Length 164490;

Best Local Similarity 86.8%; Pred. No. 1e-176;

Matches 957; Conservative 0; Mismatches 141; Indels 4; Gaps 2;

QY 725 AGGGAATGTCAGCAATCAGTTAAACCGCAACAGAGCTGAAAGGTTCTAAAGATCCTT 784

Db 5920 AGGAATGTCAACATTCATTAACCGCAACAGCTGAAAGGTTCTAAAGATCCTT 5861

QY 785 AGTTCCTCAGAGGATGCGCCGCCACTAGGAGAGCTGCGCTATATCATCTCAGAGTCC 844

Db 5860 AGTTCCTCAGAGACACGCGCCCATCTCTAGGAGAAATGTGTTTATCGTTCTGAAGCCC 5801

QY 845 CAACTCAGACAGATCTAGGAGACCTTTGCTGGATCTGGTATATCAGCATTTGTTC 904

Db 5800 CAACTCAGCGGATGTAGGAGACCTCTCTGGATCGGCATTCGACATTTGCTCA 5741

QY 905 AGCTCACCTTAGCAACGAGAGAGAGTGAACCCGCCACACACCTCAAGTAAGAGTGT 964

Db 5740 AGCTCACCAAGCAATGGACAGAGAGTGAACCCGCCACACCTCAAGTAAGAGTGT 5681

QY 965 TACTCCTCCACCTCCAGAGCGCCAGCTCCCGCTCCAGAGGTCAACTCCACCTCC 1024

Db 5680 TACTCCTCCACCTCCAGAGCGCCAGACCCCGCTCCCGGAGGACCACTCCCGCTCC 5621

QY 1025 CCCTTTCATGGGAACCAAACTCTCAAACAAAGCGCTATTCTGGAACATGGAATACGTAAT 1084

Db 5620 CCCTTTCATGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGCGGAACATGGAATACGTAAT 5561

QY 1085 CTCCGAAATCTCTCTGTCACCTGGGCGATGCAAGAGGGCTATCTCCACCACTCT 1144
 Db 5560 CTCCGAAATCTCTCTGTCACCTGGGCGATGCAAGAGGGCTATCTCCACCACTCT 5501
 QY 1145 CAACACTTCCGCAATCTCTTAATCAAGGACAGAGAGCATAGTCTGTTCTCTGT 1204
 Db 5500 TACCACTTCTCTCAATGAAATCTCTGAGGCTGAGGAGGCAATAGTCTGTTCTCA 5441
 QY 1205 TGGCAGACCAACATCATCAGCTGAGCTTCTAGCAAAATTTAACTTCCATCAGGAGACC 1264
 Db 5440 TGGTAGAACCAACATCATCAGCTGAGCTTCTAGCAAAATTTAACTTCCATCAGGAGACC 5381
 QY 1265 TGGAAATGAGAAATGTAATGCAAAATGTAATGCAAAATTTAACTTCCATCAGGAGACC 1324
 Db 5380 TGGAAATGAGAAATGTAATGCAAAATGTAATGCAAAATTTAACTTCCATCAGGAGACC 5324
 QY 1325 TGGCACTGTAATGCGGAGCAGCACTTCCATATCTCTGACAGAGCATAGTGGACAAG 1384
 Db 5323 TGGTCTGTAATGCGGAGCAGCACTTCCATATCTCTGACAGGCTAATGGACAAG 5264
 QY 1385 CCTTCTGCTTTTACAAACAGGGGATCTGCTGCTCTCTGCTCATATATACAAATGGAAT 1444
 Db 5263 CCTTCTGCTTTTACAAACAGGGGATCTGCTGCTCTCTGCTCATATATACAAATGGAAT 5204
 QY 1445 TCCTCAGTCTATGATGTTGCCAAACAGAAATAGTCAATACATGGAATATATACAAAT 1504
 Db 5203 TCCTCAGTCTATGATGTTGCCAAACAGAAATAGTCAATACATGGAATATATACAAAT 5144
 QY 1505 TGTACCTGGAGTCAAAACAAATTTGGCTCAGTCAATCTTCTGCTCAGGCGGATCATCCCC 1564
 Db 5143 TGTACCTGGAGTCAAAACAGGCTGCGGCGGATCATCCCC 5084
 QY 1565 GAGCAGTGGGATGAATCCCTACATGCGCACTTCAATACATGGAATATATACAAAT 1624
 Db 5083 GAGCAGTGGGATGAATCCCTACATGCGCACTTCAATACATGGAATATATACAAAT 5024
 QY 1625 TAATAACCAATAGGAAATAGAGCAAGTCACTCTCTGCTTCTGCTCAGGCTTCTGCTTAC 1684
 Db 5023 TAATAACCAATAGGAAATAGAGCAAGTCACTCTCTGCTTCTGCTCAGGCTTCTGCTTAC 4964
 QY 1685 AGTCACTGCAATACACAGCTCTCTTATCAACAGGCTGTGAAAGTATGCGGTATTTAA 1744
 Db 4963 AGTCACTGCAATACACAGGCTCTCTTATCAACAGGCTGTGAAAGTATGCGGTATTTAA 4904
 QY 1745 ACCAGAGTACAGCTGCTTAGACCTTACACACCTTCTTGGATACACAGGCGCAATCA 1804
 Db 4903 ACCAGAGTACAGCTGCTTAGACCTTACACACCTTCTTGGATACACAGGCGCAATCA 4845
 QY 1805 AACTGTTCAACCCAGCTCCTTTT 1826
 Db 4844 GACTGTTCAACCCAGCTCCTTTT 4823

RESULT 11
 E24616
 LOCUS 795 bp DNA linear PAT 18-JUN-2001
 DEFINITION warts protein, polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein.
 ACCESSION E24616
 VERSION E24616.1 GI:13024649
 KEYWORDS JP 1999089580-A/4.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 795)
 AUTHORS Toshiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
 TITLE warts protein, polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein
 JOURNAL Patent: JP 1999089580-A 4 06-APR-1999;
 COMMENT SUMITOMO ELECTRIC IND LTD
 OS Unidentified
 PN JP 1999089580-A/4

PD 06-APR-1999
 PF 24-SEP-1997 JP 1997258689
 PR TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
 YASUYUKI NISHIYAMA,
 PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
 G01N33/533, C12P21/02,
 PC C12N15/09, C12R1/91, (C12P21/02, C12R1/19), C12N15/00,
 (C12N15/00, PC C12R1/91)
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..795
 Location/Qualifiers
 1..795
 /organism="Unidentified"
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 224 a 228 c 160 g 183 t
 ORIGIN
 Query Match 19.8%; Score 788.6; DB 6; Length 795;
 Best Local Similarity 99.5%; Pred. No. 1.5e-163;
 Matches 791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 786 GTTCTCAGAGCATGGCCGCCCTAGAGAGAGTGTGGCTATCATCTGAGAGTCCC 845
 Db 1 GTTCTCAGAGCATGGCCGCCCTAGAGAGAGTGTGGCTATCATCTGAGAGTCCC 60
 QY 846 AACTCACAGACAGATGTAGGAAGACCTTTGTCTGATCTGGTATATCAGCATTTTCTCAA 905
 Db 61 AACTCACAGACAGATGTAGGAAGACCTTTGTCTGATCTGGTATATCAGCATTTTCTCAA 120
 QY 906 GCTCACCCTAGCAACGAGACAGAGTGAACCCGCCACCCACCTCAAGTAAAGAGTGT 965
 Db 121 GCTCACCCTAGCAACGAGACAGAGTGAACCCGCCACCCACCTCAAGTAAAGAGTGT 180
 QY 966 ACTCCTCCACCACTCCAGAGGCGCAGACTCCCTCCCAAGAGGTACAACTCCACCTCCC 1025
 Db 181 ACTCCTCCACCACTCCAGAGGCGCAGACTCCCTCCCAAGAGGTACAACTCCACCTCCC 240
 QY 1026 CTTTCATGGGAACCAACTCTCAAAACAAAGGCTATTCTGGAACATGGAATACGAATC 1085
 Db 241 CTTTCATGGGAACCAACTCTCAAAACAAAGGCTATTCTGGAACATGGAATACGAATC 300
 QY 1086 TCCGAAATCTCTCTGTCCTCCACCTGGGCGATGGCAAGAGGCTATCTCCACCACTCTC 1145
 Db 301 TCCGAAATCTCTCTGTCCTCCACCTGGGCGATGGCAAGAGGCTATCTCCACCACTCTC 360
 QY 1146 AACACTTCCCTCATGAATCTCTTAATCAAGACAGAGAGCATAGTCTGTTCTCTCTGT 1205
 Db 361 AACACTTCCCTCATGAATCTCTTAATCAAGACAGAGAGCATAGTCTGTTCTCTCTGT 420
 QY 1206 GGCAGACCAACCAATCATCATGACAGATTTCTAGCAAAATTTAACTTCCATCAGGAGACCT 1265
 Db 421 GGCAGACCAACCAATCATCATGACAGATTTCTAGCAAAATTTAACTTCCATCAGGAGACCT 480
 QY 1266 GGAATGCAAGATGTTACTGGCAAAACTGATTTATGATACACCAAAATTTGTTCTCTCTCT 1325
 Db 481 GGAATGCAAGATGTTACTGGCAAAACTGATTTATGATACACCAAAATTTGTTCTCTCTCT 540
 QY 1326 GGCAGTGTGAATCGGCGACCCACCTCCATATCTCTGACAGAGGCTAATGGACAAGC 1385
 Db 541 GGCAGTGTGAATCGGCGACCCACCTCCATATCTCTGACAGAGGCTAATGGACAAGC 600
 QY 1386 CTTTCTGCTTTTACAAACAGGGGATCTGCTGCTCTCTCTCATATACAAATGGAAGTATT 1445
 Db 601 CTTTCTGCTTTTACAAACAGGGGATCTGCTGCTCTCTCATATACAAATGGAAGTATT 660
 QY 1446 CTTCACTATGATGGTGGCAACAAAGAAATAGTCAATAACATGGAAGTATATATAACATTAGT 1505
 Db 661 CTTCACTATGATGGTGGCAACAAAGAAATAGTCAATAACATGGAAGTATATATAACATTAGT 720

QY 1506 GTACCTGGAGTGAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCAGTCATCCCG 1565
 DB 721 GTACCTGGAGTGAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCAGTCATCCCG 780
 QY 1566 AGCAGTGGGCATGAA 1580
 DB 781 AGCAGTGGGCATGAA 795
 RESULT 12
 LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 7 from patent US 6359193.
 ACCESSION AR201457
 VERSION AR201457.1 GI:20252345
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3155)
 AUTHORS Xu, T., Rao, W., Wang, W., Zhang, S. and Yu, W.
 TITLE Nucleotide sequences of lats genes
 JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
 FEATURES Location/Qualifiers
 source 1..3155
 BASE COUNT 751 a 924 c 586 t
 ORIGIN
 Query Match 19.1%; Score 760; DB 6; Length 3155;
 Best Local Similarity 63.7%; Pred. No. 3.2e-157;
 Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;
 QY 1553 CACTCTGCTAATCTCAGCCCTTCTGCTACACAGTCACTGCAATTCACACAGCTCTCTATT 1712
 DB 985 CAACCTGAGCCCTCACTGCGCCGCCCAACACAGCGTCAACCGCTGAGCGCCGACACATC 1044
 QY 1713 CAACAGCTCTGAAAGTATGCTGTATTAACACAGAGTACAGACTGCTTTAGCACCT 1772
 DB 1045 CTTACCTCTGAAAGCGTGTGCTGCTGCGCCGCCAGCCAGACGCGTGGGCC 1104
 QY 1773 ACACAGCTCTTTGGATPACACAGCCAAATTCAAACTGTTCAACCCAGTCTCTTTCTCTGAG 1832
 DB 1105 TCGCACCCCGCTGGTGGTGGCTGCGCCCAACAGCACCTGCCACTGAGAGCTGGAGACGAAG 1164
 QY 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTTCTGAAGTCCAAACTATCAAGGA 1892
 DB 1165 GAGGCGACGCGCCGCCACACCCCGCTGGATGTGGACTATGGCGCTCCGAGCGGAGTGC 1224
 QY 1893 CCACACACCCCTACCCAAACATCTGCTGCACCAACCCATCTGTTCTCTCCATACGAG 1952
 DB 1225 CCACCGCTCGTATCCAAAGCACTGCTGCTGCCAGTAACTGAGCAGTACAGCGTG 1284
 QY 1953 TCAATCAGTAAAGCTAGCAAGAGGATPACCCAGCTTGGCCCAAGAGATGAGAGTGAA 2012
 DB 1285 GACCTGGACAGCTGTGCACAGTGTGCAGAGTCTGCGAGGGGCACTGATCTAGAC 1344
 QY 2013 AAGAGTTA-----TGAATGTTTCATAGTGGGATGAAGAAAGAAACAG 2057
 DB 1345 GGGAGTGACAGAGCCCAAGAGTGGCAAGGGAGAGCAAAAGCTGGCAGAGCAAAAGAG 1404
 QY 2058 ATTAACAATCTCACCTATTACTGTTAGGAAACAAAGAAAGATGAAGAGCGAAGGAAATCT 2117
 DB 1405 ATTACAGCTCTCCCGTGGCTGTCGCGCAAGNATAGCAGATGAAGAGAGAGAGTCT 1464
 QY 2118 CGTATTCAAAGTATTCTCCTCAAGCAATTTAAATCTTTTATGGAGCAACATGTAGAAAT 2177
 DB 1465 CGCATCAAGAGTATTCTCCCTTATGCTTCAATTTCTTATGGAGCAACAGCTGGAGAAAT 1524
 QY 2178 GTACTCAATCTCATCAGCAGCTCTACATGTTAAACAAATTTAGAGAAATGAATGATG 2237
 DB 1525 GTCATCAAAACCTTACCAGCAGAGGTCAGCGGAGGCTACAGCTGGAGCAGGAATGGCC 1584

QY 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGGCCAAAAGAA 2297
 DB 1585 AAAGCTGGGCTCTGTGAGCCGAGCAGAGATGAGGAAGATCTCTACAGAAGGAG 1644
 QY 2298 TCTAATTACATCCGCTTTAAAGGGCTAAATGGACAGTCTATGTTGTGAAGATAAG 2357
 DB 1645 TCTAATACACCGCTGAAGAGGCCAAGATGGACAGTCCATGTTGTGAAAATCAAG 1704
 QY 2358 ACACCTAGGATAGGAGCATTTGCTGAAGTCTGTAGCAAGAAAGATAGATCTAAGGCT 2417
 DB 1705 ACTCTAGGATCGGTGCTTTGGGAAGTGTCCCTGCTGTGAAGTGGACACTCACGCT 1764
 QY 2418 TTGTATGCAACAAAACCTCTCGAAAGAAAGATGTTCTTCTCGAAATCAAGTGCCTAT 2477
 DB 1765 CTGTAGCCCATGAAGACTCTCAGGAAGAGGATGCTCTCAACCGGAATCAAGTGGCCAT 1824
 QY 2478 GTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACAATGAATGGTAGTCTGCTATAT 2537
 DB 1825 GTCAAGGCTGAGAGGAGACATCTGCTGAAGCAGACAATGAGTGGGTGCTAAACTCTAC 1884
 QY 2538 TATTCATTCCAAGATAAGACAAATTTATATCTTGTAAATGAGACTACATTCCTGGGGTAT 2597
 DB 1885 TACTCTCTCCAGACAGGACACCTGTACTTTGTGATGAGACTACATACAGCGGGAT 1944
 QY 2598 ATGATCAGCTTATTAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGATCTACATA 2657
 DB 1945 ATGATCAGCTTCTGATCAGGATGAGGCTCTCCCTGAGCACCTGGCCGCTTCTACATT 2004
 QY 2658 GCAGAACTTACCTGCTGAGTGAAGTGTTCATAAATGGTGTTCATAGAGATAT 2717
 DB 2005 GCAGAGTTGACCTGGCCATTTGAAGTGTCCCAAGATGGGCTTTATCCACCGGACATC 2064
 QY 2718 AAACCTGATAAATTTTGTATGATCGTGTATGATGTCATATTAATTTGACTTTGGCCCTC 2777
 DB 2065 AAGCTTGACAACTATCTACGACCTGGATGGTATGTCATATTAAGCTGACAGATTTGGCCCTC 2124
 QY 2778 TGCACTGGCTTACAGTGGACACAGATTTCTAAGTACTATCAGAGTGTGACCATCCAGG 2837
 DB 2125 TGCACTGGATTTCAAGTGGATCAATTCACAGTACAGAAAGGAAACCAATGAGA 2184
 QY 2838 CAAGATAGATGATTTCAAGTGAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 2897
 DB 2185 CAGGACAGATGAGGCGCGGTGACCTCTGGGAGATGTTTCCAACTGTGCTGTGGAGAC 2244
 QY 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGACAGCAGAGAGTGCCTGGCAGATCT 2957
 DB 2245 AGCTTAAAGACCTTGAGCAGAGAGGGCGCAGACAGCAGAGTGCCTGGCAGATCT 2304
 QY 2958 TTGCTGGGACTCCCAATTTATTTGACCTGAAGTGTGCTAGCAAGAGATACACAG 3017
 DB 2305 CTGCTGGGACACAAATTTACATCGCTCCGAGGTGCTTCTCCGCAAGGGTACACGAG 2364
 QY 3018 TTCTGTGATTTGGAGTGTGGTGTATTTCTTTTGAATGTTGGTGGGACAACTCTCT 3077
 DB 2365 CTCTGTGACTGGTGGAGCGTGGTGTGATTTCTCTTTGAGATGCTGTTGGGACGCGCT 2424
 QY 3078 TTCTTGGCACAACACCAATTTAGAAACACAAATGAAGSTTATCAACTGGGCAACATCTCT 3137
 DB 2425 TTCTTGGCCCCCAACCCCAACAGACAGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTG 2484
 QY 3138 CACATTCACCAACAGCTAACTCAGTCTCCTGAGCTTCTGATCTTATTTAATTAACCTTGC 3197
 DB 2485 CATATCTCTACGAGTGGAGCTCAGCGTGGGCGCCGAGACCTCATCAGAGAGTGTGC 2544
 QY 3198 CGAGGACCCCAAGATCGCTTAGCAAGAAATGGTCTGCTGATAAATAAAGCTCATCATTT 3257
 DB 2545 TGGCGGCTGACTGCCCTGGGCGAGGATGGGCGAGATGACCTCAAGGACACCCGCTC 2604
 QY 3258 TTTAAACAAATGACTTCTCCAGTAGCTGAGACAGCAGTCTGCTTCATACATCTCTAAA 3317
 DB 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGCAGGCTGACACCTACGCTCCCAAC 2664

polyA_site		/gene="MmLAT52"		3440		EGAAAGCQPVVV		SGSAAKAWDTLASPSKPHBEAFYEYTERFRFDNGYPTFRCPKPSPEASADPGDADL		CRILRGDADDLKAHPNTIDFSRIRKQPPAYPIISHPMDTISNDFPDVDESPPHEA		VGVILFEMLVGQPPFLAPTPETQLKVINWESTLHIPTQVRLSAEARDLITKLCCAA	
BASE COUNT	861 a	992 c	963 g	644 t									
ORIGIN													
Query Match	19.1%;	Score 760;	DB 10;	Length 3460;									
Best Local Similarity	63.7%;	Pred. No. 3.2e-157;											
Matches 1197;	Conservative	0;	Mismatches 665;	Indels 18;	Gaps 2;								
QY	1653	CAC	TCTGCTAATTC	TCTAGCC	TTTGCTACACAGTCACTGC	CAATACACGAGTCCTTATT	1712						
Db	1304	CAACCTGAGCCCT	CAC	TGCCCGCC	CCCAACACGCTACCCGCGTGACGGCGGCACACATC	1363							
QY	1713	CAACAGCCTGT	GAAAGCTAT	TGCTGTTATT	AAACACGAGCTACAGACTGCTTTTAGCACCT	1772							
Db	1364	CTTCACCTGT	GAAGACG	TGCTGCTGCTGGGCCC	GAGCCCCACAGACGCGTGGGCCCC	1423							
QY	1773	ACACACCTTT	TGTGGATPAC	CACAGCCAA	TCAAAC	TGTTCAACCCAGTCCTTTTCTCGAG	1832						
Db	1424	TGCGACCCCGC	TGGTGGT	TGCCCCAC	ACGACCTGCCACTGAGAGCCCTGGAGACCAAG	1483							
QY	1833	GGACCCGCTT	CAATGTG	ACTGTGTG	TATCCACCTTGTCTG	NAAGCTCCAAACTATCAAGGA	1892						
Db	1484	GAGGCGAGCG	AGGCC	CCACACCCGCTGG	ATGTGGACTATGGCGCTCCGAGCGCAGGTGC	1543							
QY	1893	CCACACACG	CCCTACCC	CAAAACAT	CTGCTGCACCAAAACCCATCTGTTCTCTCCATACGAG	1952							
Db	1544	CCACCGCTTC	GTACCC	CAAAAGCACTT	CTGCTGCCAGTAAGTCTGACGAGTACAGCGTG	1603							
QY	1953	TCAATCAGT	AAGCCTACA	AGAGGATC	AGCAGCTTGCCCAAGGAAGATGAGAGTGAA	2012							
Db	1604	GACCTGCAC	AGCCCTGTG	CACCACTGTG	CGACAGAGTCTGCGAGGGGCGCACTTGAGCAAGAC	1663							
QY	2013	AAGAGTTA	-----	-----	TGAAATGTTTCATAGTGGGGATAAGAAAGAAACACAG	2057							
Db	1664	AGNATCA	AGAGCC	CAAAAGGTG	CCAGGGAGACAAGCTGGCAGAGACAAAGACAG	1723							
QY	2058	ATTACA	ACTTCA	CTATTACT	GTGTAGGAAACAAAGAAAGATGAAGAGCGAAGGGAATCT	2117							
Db	1724	ATT	CAGACCT	CCCGG	TGCTGCTCCGCAAGAAATAGCAGAGATGAAGAGAGAGAGAGCTCT	1783							
QY	2118	CGTAT	TCAAAGT	TATTTCT	CTCTCAAGCATTTAAATCTCTTTATGGAGCAACATGTAGAAAT	2177							

Db	1784	CGCATCAAGAGTTTACTCCCTTATGCCCTTCAAAATCTTTCATGGAGCAACACGTGGAGAA	1843
Qy	2178	GTACTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAAATTTAGAGAATGAATCATG	2237
Db	1844	GTTCATCAAAACCTTACCAGCAGAAGGTGAGCGGAGGCTCAGCTGGAGCAGGAAGTGGCC	1903
Qy	2238	CGGGTTGGATATCTCAAGATGCCAGGATCAATAGAGAAGATGCTTTGCCAAAAAGAA	2297
Db	1904	AAAGCTGGGCTCTGTGAGCGGAGCAGGACGATGAGGAAGATCCTCTACCAGAGGAG	1963
Qy	2298	TCTAATTTACATCCGCTTTAAAAGGCTTAAATGGCAAGTCTATGTTGTGAAGATAAAG	2357
Db	1964	TCTAACTACAACCGGCTGAAGAGGGCCAAAGTGCACAGTCCATGTTGTGAAAAATCAAG	2023
Qy	2358	ACACTAGGAATAGAGACATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGATACTTAGGCT	2417
Db	2024	ACTCTAGGCATCGTGCTTTGGGAAAGTGTGCTCGCTTGAAGCTGGACACTCAGGCT	2083
Qy	2418	TTGTATGCACAAAACACTCTTCAAGAAAGATGTTCTTCTTGAATCAAGTCGGCTCAT	2477
Db	2084	CTGTACGCCATGAAGACTCTCAGGAAGAAGATGTCTGAACGGGAATCAAGTGGCCAT	2143
Qy	2478	GTTAAGGCTCAGAGAGATATCCTTGGCTGAAGCTTGACAAATGAATGGGTAGTTCGTCATAT	2537
Db	2144	GTCAAGGCTCAGAGGGACATCCTTGGCTTGAAGCAGCAATGAGTGGGTGGTCAAACTTCAC	2203

QY 2538 TATTTCATCCAGATAGACATTTATTTACTTTGTAATGACTACATTCCTCGGGTGAT 2597
 DB 2204 TACTCTTCCAGACAGGACACCGCTGTACTTTGTGATGGACTACATACAGCGGGGAT 2263
 QY 2598 ATGATGAGCCTATTAAATAGAAATGATCTTTCCAGAAAGTGTGCAGATTTACATA 2657
 DB 2264 ATGATGAGCCTGTGATGAGGAGGCTTCCCTGAGCACCTGCCCGCTTACAT 2323
 QY 2658 GCAGAACTTACCTGTGAGTTGAAATGTTGATATAAATGGGTTTATTATAGAGATAT 2717
 DB 2324 GCAGAGTTGAGCTTGGCCATTAAGAGTTGCTCAAGATGGCTTTATCCACGGGACATC 2383
 QY 2718 AAACCTGATATATTTGATGATGCTGATGCTCATATAAATGACTGACTTTGGCCTC 2777
 DB 2384 AAGCCTGACAACTACTCATCGACCTGGATGGTCAATTAAGCTGACATTTTGGCCTC 2443
 QY 2778 TGCACTGGCTTCAGATGACACACGATTTCTAAGTACTATCAGAGTGTGACCATCCACGG 2837
 DB 2444 TGCACTGGATCAGGTGGACTCACAATTCACAGTACTACAGAAAGGAACACATGAGA 2503
 QY 2838 CAAGATAGCATGATTTCAATGATGATGAGTGGGATCCCTCAAGCTGTGATGTGAGAC 2897
 DB 2504 CAGGACAGCATGGAGCGCGGTGACCTCTGGGACGATGTTTCCAACTGTGCTGTGGAGC 2563
 QY 2898 AGACTGAAGCCATTAGAGCGGAGAGTGCACGCCAGCAGCGATGCTCTAGCACATTC 2957
 DB 2564 AGTTAAGACCTTGGAGCAGAGGGCCGACAGACACACAGAGGTGCTGGGACATTC 2623
 QY 2958 TTGTTGGGACTCCCAATTTATTTGACCTGAAATGTTGCTAGCAACAGGATACACAGC 3017
 DB 2624 CTTCTCGGACACCAATTTACATCGCTCCGAGGTGCTTCTCCGCAAGGGTACAGCGAG 2683
 QY 3018 TTGTGCTATTTGGGAGTGTGTTGTTATTTCTTTTGAATGTTGTTGGGACACCTCT 3077
 DB 2684 CTCTGTGACTGGTGGAGCGGTGCTGATTTCTTTGAGATGCTGTTGGGACCGCCT 2743
 QY 3078 TTCTTGGCACAACACCATTTAGAACACAAATGAAGTTTATCACTGGCAACATCTCTT 3137
 DB 2744 TTCTTGGCCCCCACCACACAGAGACGAGCTGAAGTGTATCAACTGGGAGAGCAGCTG 2803
 QY 3138 CACATTCACCAAGCTAACTCAGTCTGAGCTTCTGATCTTATTATTAACCTTGC 3197
 DB 2804 CATATCTCTAGCGAGTGAGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCT 2863
 QY 3198 CGAGGACCGAAGATCGTTAGGCAAGATGGTCTGATGAAATATAAAGCTCATCTCAAT 3257
 DB 2864 TGCGGGGTGACTCCGCTGGGAGGATGGGAGATGACCTCAAGCAGCACCCCGTTC 2923
 QY 3258 TTTAAACAAATTTGACTTCTCAGTGACTGAGACAGAGTCTGCTTCAATTCCTTAA 3317
 DB 2924 TTCAACACCATCGACTTTTCCCGTGACATCGAAAGCAGCTGCACCTACGTCGCCACC 2983
 QY 3318 ATCACACACCAACAGATACATCAATTTGATCTCTGATCTCTGATAAATTTATGAGT 3377
 DB 2984 ATCAGCACCACCATTTGACACCTTCAATTTTACCGGTGGATGAAGAAGCCCTGGCAC 3043
 QY 3378 GATGATAACGAGGAAGAAATGTAATGACACTCTCAATGGATGGTATAAATAATGGAAG 3437
 DB 3044 GAGGCCAGC---GGAGAGCGCCCAAGCCCTGGGACAGCTGGGCTCCCGCCAGCAAG 3100
 QY 3438 CATCTGAAACATGATTTCTATGATTTACCTTCGAGAGTTTGTGATGACAAATGGCTAC 3497
 DB 3101 CATCAGAGCAGCGCTTTATGATTTACCTTCCGAGGTTTCTTCGATGACAAAGGCTAT 3160
 QY 3498 CCATATAATTTCCGAAGCC 3517
 DB 3161 CCTTCCGCTGCCGAAGCC 3180

RESULT 14
 E38226
 LOCUS Human tumor regulatory gene
 DEFINITION Human tumor regulatory gene
 5486 bp DNA 1linear PAT 31-JAN-2002

E38226
 E38226.1 GI:18626935
 JP 2000210086-A/1.
 Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5486)
 Koda, J., Kono, K. and N.Z.F.
 Human tumor regulatory gene.
 Patent: JP 2000210086-A 1 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 OS Homo sapiens (human)
 PN JP 2000210086-A/1
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
 C12N15/09, A61K31/00, A61K38/27, C07K14/47, C07K14/525, C12Q3/68, PC
 G01N33/50,
 PC G01N33/50, G01N33/574, G01N33/577//C12P21/08, (C12N15/09, PC
 C12R1-91), C12N15/00,
 PC A61K37/36, (C12N15/00, C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT source 1. 5486
 FT /organism="Homo sapiens"
 FT /db_xref="taxon:9606"
 BASE COUNT 1378 a 1452 c 1390 g 1266 t
 ORIGIN
 Query Match 18.7%; Score 744.4; DB-6; Length 5486;
 Best Local Similarity 66.3%; Pred. No. 9.1e-154;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;
 QY 1872 GAAGTCCAAACTATCAAGGACACACACACCTACCCACCAAAACATCTGCT----GCACC 1926
 DB 1905 GGAGGCCAGACCGGAGTGCCTCGCTACCCGAGCCTGCTGTGGCGAGC 1964
 QY 1927 AAACCCATCTGTTCCATAGGATCAATAGTAACTAGCAAGAGGATCAG---- 1982
 DB 1965 AGTCGGAGCAGTACGACCTGGACACCTGTGGCAGGATGGAGCAGAGCTCCGTGGG 2024
 QY 1983 ---CCAAAGTTGCCCAAGGAAGATGAGAGTGAAGAGAGTTATGAAATGTTAGTAGTGG 2039
 DB 2025 GGCCCCACGAGCGCGGCGGCGGCGGACAGAGCGCCAAAGCGCCAAAGGACAAAGC 2084
 QY 2040 GATAAAGAAAGAAACAGATTACACTTACCTTACTTACTTACTTACTTACTTACTTACTT 2099
 DB 2085 GGAAGGATAAAAGCAGATTACAGCTTCTCCGCTTCCCGTCCGCAAAACACAGAGAC 2144
 QY 2100 GAAGCGAAGGAATCTCGTATTCAAAGTTATTTCTTCTCAAGCATTTAAATCTTTATG 2159
 DB 2145 GAAGAGAGAGAGTACCCATCAAGACTACTCGCCATACGCTTTAAGTTCTTCATG 2204
 QY 2160 GAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGAGGCTTACATCGTAAACAA 2219
 DB 2205 GAGCAGCAGCTGAGAAATGTATCAAAACCTTACAGCAAGGTTAACCGAGGCTGCAG 2264
 QY 2220 TTAGAGATCAATGATGCGGTTGGATTATCTCAAGATCCCGAGGATCAATAGAGAAAG 2279
 DB 2265 CTGAGCAAGAAATGGCCAAAGCTGGACTCTGTAGCTGAGCAGGAGAGATGCGGAAG 2324
 QY 2280 ATGCTTTGCCAAAAGAAATCTAATTTACATCCGCTTTAAAGGGCTAAATGGACAGTCT 2339
 DB 2325 ATCTCTACCAAGAAAGTCTAATTTACACAGGTTAAAGAGGGCCAAAGTGGACAGTCT 2384
 QY 2340 ATGTTTGTGAAGATAAGACACTAGGATAGGAGCATTTTGTGAGTCTCTCTAGCAAGA 2399
 DB 2385 ATGTTTGTGAAGATAAGACACTAGGATAGGAGCATTTTGTGAGTCTCTCTAGCAAGA 2444

QY 2400 AAAGTAGATACAAAGCTTTTGTATGCAACAAACAACTCTTCGAAGAAAGATGTCTTCTT 2459
 Db 2445 AAGGTGGACATCACGCCCTGTACGCCATGAAGACCCCTAAGGAAAGATGTCTCTGAAC 2504
 QY 2460 CGAATCAAGTCTCATCTTAAAGCTGAGAGAGATATCTCGCTGAAGCTGACAAATGA 2519
 Db 2505 CGAATCAAGTCTCATCTTAAAGCTGAGAGAGATATCTCGCTGAAGCTGACAAATGA 2564
 QY 2520 TGGGTAGTCTCTATATATATTCATTCACCAAGATAGGACAATTTATATCTTGTATGGAC 2579
 Db 2565 TGGGTGGTCAACTCTACTACTCTTCCCAAGACAAGACAGCTGTACTTTGTATGGAC 2624
 QY 2580 TACATCTCTGGGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2639
 Db 2625 TACATCTCTGGGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2684
 QY 2640 CTGGACAGATCTACATACAGACATCTACCTGTGAGATGATGATGATGATGATGATGAT 2699
 Db 2685 CTGGCCGGTCTTACATCGCAGAGCTGACTTTGGCCATTTGAGAGTGTCCACAGATGGC 2744
 QY 2700 TTTATTCATAGAGATTAACCTGATATATTTTGAATGATGATGATGATGATGATGATGAT 2759
 Db 2745 TTTATTCATAGAGATTAACCTGATATATTTTGAATGATGATGATGATGATGATGATGAT 2804
 QY 2760 TTGACTGACTTTGGCTCTGCACTTTCAGATGATGATGATGATGATGATGATGATGATGAT 2819
 Db 2805 CTCACAGATTTTGGCTCTGCACTTTCAGATGATGATGATGATGATGATGATGATGATGAT 2864
 QY 2820 AGTGTGACATCCACGGCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2879
 Db 2865 AAGGGAGAGCTGTGACAGAGGACAGATGATGATGATGATGATGATGATGATGATGATGAT 2924
 QY 2880 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2939
 Db 2925 AACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2984
 QY 2940 CGATGTCTAGACATCTTTTGGTGGGACTCCCAATATATATGATGATGATGATGATGATGAT 2999
 Db 2985 AGGTGCTGGACATTCATCTGTGGGACTCCCAATATATATGATGATGATGATGATGATGAT 3044
 QY 3000 CGAAGAGATACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3059
 Db 3045 CGAAGAGATACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3104
 QY 3060 TTGGTGGGACACCTCTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGAT 3119
 Db 3105 CTGGTGGGACACCTCTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGAT 3164
 QY 3120 AACTGGCAACATCTTTCACATTCACCAAGCTAACTCAGTCTGATGATGATGATGATGATGAT 3179
 Db 3165 AACTGGCAACATCTTTCACATTCACCAAGCTAACTCAGTCTGATGATGATGATGATGATGAT 3224
 QY 3180 CTTATATTAATTTTGGAGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3239
 Db 3225 CTTATATTAATTTTGGAGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3284
 QY 3240 ATAAAGCTATCTCTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGATGAT 3299
 Db 3285 CTTATATTAATTTTGGAGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3344
 QY 3300 GCTTCATATCTTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGATGATGAT 3359
 Db 3345 GCTTCATATCTTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGATGATGAT 3404
 QY 3360 CTTATATTAATTTTGGAGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGAT 3419
 Db 3405 GCTTCATATCTTCTTCTTGGCACAACACATGATGATGATGATGATGATGATGATGATGATGAT 3461
 QY 3420 TGGTATAAATGGAAGATCTCTGACATGATGATGATGATGATGATGATGATGATGATGATGAT 3479
 Db 3462 ACCTGCCCAATTAACAGCATCTGACGACGATTTTACGATGATGATGATGATGATGATGATGAT 3521

QY 3480 TTTGATGACAATGGCTTACCCATATATATTCGAGCCCTATTGAATATGAATACATTAAT 3539
 Db 3522 TTTGATGACAATGGCTTACCCCTTTCGATGCCCAAGACCTTCAGGAGCAGAAGCTTCACAG 3581
 QY 3540 TCATAGAGGCTCAGA 3553
 Db 3582 GCTGAGAGCTCAGA 3595

RESULT 15
 E38227
 LOCUS Human tumor regulatory gene.
 DEFINITION Human tumor regulatory gene.
 ACCESSION E38227
 VERSION E38227.1 GI:18626936
 KEYWORDS JP 2000210086-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE
 1 (bases 1 to 5486)
 Koga,J., Kono,K. and N.Z.F.
 Human tumor regulatory gene
 Patent: JP 2000210086-A 2 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 OS Homo sapiens (human)
 PN JP 2000210086-A/2
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR
 PI JUNICHI KOGA,KEIKO KONO,ZOROTARYO FUJIOH N PC
 C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68, PC
 G01N33/50,
 PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09, PC
 C12R1:91),C12N15/00,
 PC A61K37/36,(C12N15/00,C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT source 1..5486
 FT /organism='Homo sapiens (human)'.
 FEATURES
 source
 1..5486
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 1380 a 1451 c 1388 g 1267 t
 ORIGIN

Query Match 18.7%; Score 744.4; DB 6; Length 5486;
 Best Local Similarity 66.3%; Pred. No. 9.1e-154;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;

QY 1872 GAAGTCCCAAACTATCAAGGACCCACCCCTACCCCAAAACATCTGCT-----GCACC 1926
 Db 1905 GGAGGCCAGACGGAGGTCGCCCTTCCGCTTCCGCAAGACCTGCTGCTGCGCAGC 1964
 QY 1927 AAAACCCATCTCTTCTCCATACGATCAATCAGTAAGCTTAGCAAGAGGATCAG----- 1982
 Db 1965 AAGTCGAGCAGTACGACCTTGGACGCTGTCGCGAGGATGAGCAGAGCTTCGTCG 2024
 QY 1983 ---CCAAGCTTCCCAAGGAAGATGAGAGTGAAGAGATGATGAAATGTTGATAGTGG 2039
 Db 2025 GGCCCAACAGCCGAGGCGGACAGAGCCGCAAAAGCCCAAGGGGACAAAGGC 2084
 QY 2040 GATAAAGAAAACAGACATTAACCTTACCTTATCTGTTAGGAAAACAGAAAGAT 2099
 Db 2085 GGAAGAGTAAAGACGATTCAGACCTTCCGCTTCCGCAAGACAGCAGAC 2144
 QY 2100 GAAGAGGAGGAATCTGATTTCAAGTATTTCTCTCAAGCATTTAAATCTTTATG 2159
 Db 2145 GAAGAGAGAGAGTACGCGATCAAGAGCTACTGCCATAGCCTTTTAAGTTCTTCATG 2204
 QY 2160 GAGCAACATGTAGAAATGTTACTCAATTCATCAGCAGCGTCTATCGTAAAAACAA 2219

Db 2205 GAGCAGACCGTGGAGATGTCTATCAAAACCTACCAGCAGAGGTTAAACCGAGGCTGCAG 2264
Qy 2220 TTACAGAAATGAATGATGCGGGTTGGATTATCTCAAGATCCCGAGATCAAAATGAGAAAG 2279
Db 2265 CTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGAGATCGCGAAG 2324
Qy 2280 ATGCTTTGCCAAAGAAATCTAATTACATCCGCTTTAAAGGGCTAAATGGCAAGTCT 2339
Db 2325 ATCTCTTACCAGAAAGAGTCTAATTACACAGGTTAAAGAGGGCCAAAGATGGACAGTCT 2384
Qy 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTTAGCAAGA 2399
Db 2385 ATGTTTGTCAAGATCAAAACCCCTGGGATCGGTGCCCTTTGGAGAGTGTGCCCTGCTGT 2444
Qy 2400 AAAGTAGATACTAAGGCTTTGTATGCAACAAAAACTCTTCGAAGAAAGATGTCTTCTT 2459
Db 2445 AAGGTGGACACTACGCCCTGTACGCCATGAAGACCCCTAAAGGAAAGAGATGCTCTGAAC 2504
Qy 2460 CGAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCCTGGCTGAAGCTGCAATGAA 2519
Db 2505 CGGAATCAGGTGGGCCACGCTCAAGCCGAGAGGACATCTGGCCGAGGACAGCAATGAG 2564
Qy 2520 TGGTAGTTCCTATATTAATTCATTCGAAGATAAGGACAAATTTATACTTTGTAATGGAC 2579
Db 2565 TGGTGGTCAAACTCTACTACTCTCTCAAGACAAAGACGCTGTACTTTTGTGATGGAC 2624
Qy 2580 TACATTCTGGGGTGATATGATGAGCCTATTAAATGAATGGGCATCTTCCAGAAAGT 2639
Db 2625 TACATCCCTGGTGGGACATGATGAGCTGCTGATCCGGATGGAGGCTTCCCTGAGCAC 2684
Qy 2640 CTGGCAGCATTTCTACATAGCAGAACTTACCTGTGCAAGTTGAAGTGTTCATAAATGGT 2699
Db 2685 CTGGCCCGGTTCTACATCGCAGAGCTGACTTTGGCCATGAGAGTGTCCCAAGATGGC 2744
Qy 2700 TTTATTCATAGAGATTAATAACCTGATTAATTTTGTATGATCGTGAATGATGATTAATA 2759
Db 2745 TTATCCACCGAGACATCAAGCCTGATAAATTTTGTATGATGATCTGGATGGTCAATTA 2804
Qy 2760 TTGACTGACTTTGGCTCTGCACCTGGCTTCAGATGGACACAGATTTCTAAGTACTATCAG 2819
Db 2805 CTCACAGATTTGGGCTCTGCACCTGGGTTGAGTGGACTCAAAATCCAAATATACCAG 2864
Qy 2820 AGTGTGACCATCCACGGCAAGATAGCATGATTAATGAATGAATGGGGGATCCCTCA 2879
Db 2865 AAAGGACCCATGTGACAGACAGCAGCATGGAGCCAGCCTCTGGGATGATGTGTCT 2924
Qy 2880 AGCTGTGATGAGAGACAGACTGAGCCATAGAGCGGAGAGCTGCACGCCAGCACAG 2939
Db 2925 AACTGTGCTGGGACAGGCTGAAGACCCCTAGAGCAGAGGGCGGGAAGCAGCACCCAG 2984
Qy 2940 CGATGCTAGCACATTTCTTGGTGGGACTCCCAATTTATTTGACCTGAAAGTGTGCTA 2999
Db 2985 AGTGCCTGGACATTTCACTGGTGGGACTCCAACTACATCGCACCGGAGGTGCTCCTC 3044
Qy 3000 CGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTCTTTTGAATG 3059
Db 3045 CGCAAGGGTACACTCACTCTGTACTGTGGTGGTGTGGAGTGATTTCTCTCGAGATG 3104
Qy 3060 TTGTTGGGACACCTCTCTTCTTGGCAAAACACCATTAAGAACACAAATGAAGTTATC 3119
Db 3105 CTGGTGGGGACGCCCTTTTGGCACCCTACTCCACAGAACCCAGCTGAAGTGATC 3164
Qy 3120 AACTGGCAAAACATCTTTTACATTTCCACCACAAAGTAACTAGTCTGAAAGCTTCTGAT 3179
Db 3165 AACTGGGAGAACGCTCCACATTTCCAGCCAGGTGAAGCTGAGCCCTGAGGCCAGGGAC 3224
Qy 3180 CTTATTATTAACCTTTGGCGAGGACCCGAGATCCCTTAGGCAAGATGGTCTGATGAA 3239
Db 3225 CTCATCACCAAGCTGTGTGCTCCGACAGACCACCCCTGGGGCGGAATGGGGCCGATGAC 3284
Qy 3240 ATAAAGGCTCATCCATTTTAAAAACAATTTGACTTCTCCAGTGACCTGACAGCAGTCT 3299
Db 3285 CTGAAGGCCACCCCTCTTTCAGGCCATGACTTCTCCAGTGACATCCGGAAGCAGCCA 3344

Qy 3300 GCTTTCATACATTCTCTAAAATCACACACCCCAACAGATACATCAAAATTTTGTCTGTGAT 3359
Db 3345 GCCCCCTAGTTCCTCCACCAATCAGCCACCCCAATGGACACCTCGAATTTCCACCCCGTAGAT 3404
Qy 3360 CTTGATAAAATTTATGGAGTGTATGATAACGAGGAGAAATGTAATGACACTCTCAATGGA 3419
Db 3405 GAAGAAAGCCCTTGGAAACGATGCCAGC---GAAGGTAGCACCACCAAGGCTTGGGACACACTC 3461
Qy 3420 TGGTATAAAATGGGAAAGCATCCTGAACATGCAATTCATGAATTTACCTTCCGAAGGTTT 3479
Db 3462-ACCTCGCCCAATAAACAGCATCTGAGCAGCATTTTACGAATTTACCTTCCGAAGGTTT 3521
Qy 3480 TTTGATGACAATGGCTACCCATTAATTTCCGAAGCCTATTGAATATGAATACATTAAT 3539
Db 3522 TTTGATGACAATGGCTACCCCTTCCGATGCCCAAGGCTTCAGGAGCAGAAAGCTTCCACAG 3581
Qy 3540 TCACAAGGCTCAGA 3553
Db 3582 GCTGAGAGCTCAGA 3595

Search completed: January 16, 2003, 18:10:29
Job time : 7470.05 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 508.163 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: 1 accttgggtgtctgggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_101002:*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3984	100.0	3984	17	AAT42118
2	3984	100.0	3984	21	AAZ51505
3	3978.8	99.9	7382	20	AAZ87396
4	2483	62.3	3213	21	AAZ51506
5	2479.8	62.2	3213	17	AAT42119
6	2422	60.8	2442	20	AAZ32981
7	1360.4	34.1	1374	20	AAZ32982
8	1233.2	31.0	1357	22	ABA08740
9	788.6	19.8	795	20	AAZ32984

10	760	19.1	3155	17	AAT42120
11	760	19.1	3155	21	AAZ51507
12	744.4	18.7	5276	20	AAZ87397
13	744.4	18.7	5486	21	AAZ51508
14	744.4	18.7	5486	21	AAZ51509
15	744.4	18.6	1961	21	AAZ51510
16	742.8	18.6	3533	22	AAZ51511
17	741.8	18.6	3533	22	AAZ51512
18	693.8	17.4	1912	23	ABK43457
19	675.8	17.0	1498	21	AAZ61158
20	655.4	16.5	2043	22	AAZ61159
21	609.8	15.3	638	21	AAZ79942
22	580.4	14.6	582	20	AAZ32983
23	478.6	12.0	1501	22	AAZ75341
24	469	11.8	5720	17	AAT42117
25	467.4	11.7	3319	23	ABL03169
26	452.8	11.4	678	21	AAZ79982
27	346.2	8.7	676	22	AAZ27179
28	346.2	8.7	676	23	ABK43776
29	269.4	6.8	1935	20	AAZ06834
30	269.4	6.8	4983	22	AAZ44629
31	254.6	6.4	3018	17	AAT31452
32	254.6	6.4	3018	24	ABK84773
33	254.6	6.4	3800	22	ABA08668
34	233	5.8	1710	21	AAZ34303
35	233	5.8	2003	21	AAZ49841
36	218.2	5.5	2001	21	AAZ36174
37	216.8	5.4	11187	23	ABL03168
38	210.6	5.3	1894	21	AAZ39567
39	209	5.2	1922	21	AAZ49426
40	186.4	4.7	734	23	ABK43709
41	186	4.7	734	22	AAZ27161
42	186	4.7	734	22	AAZ56728
43	186	4.7	734	23	ABK43985
44	180.4	4.5	2255	23	ABL29695
45	173.2	4.3	2160	22	AAI66703

ALIGNMENTS

RESULT 1	
AAZ42118	
ID	AAT42118 standard; cDNA; 3984 BP.
AC	AAT42118;
XX	
DT	31-JAN-1997 (first entry)
DE	H-lats gene encoding large tumour suppressor.
XX	
KW	Human; h-lats gene; large tumour suppressor; fetal brain;
KW	protein-serine/threonine-kinase; cell proliferation; antisense;
KW	dominant-negative; cancer; degenerative disorder; trauma;
KW	growth deficiency; therapy; antitumour; vulnary; diagnostic;
KW	transgenic plant; transgenic animal; growth; senescence; ds.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	231..3623
CDS	/*tag= a
FT	/product= Lats protein
XX	
PN	WO9630402-A1.
XX	
PD	03-OCT-1996.
XX	
PF	26-MAR-1996; 96WO-US04101.
XX	
PR	27-MAR-1995; 95US-0411111.
XX	
PA	(UYVA) UNIV YALE.

M-lats2 gene encod
Mouse Lats2 (large
Human WART2 cDNA.
DNA encoding a tum
cDNA SSI171A encod
Human cancer-inhib
DNA encoding novel
cDNA SSI171 encodi
Human protein tyro
Human colon cancer
Human warts gene f
Human TGF-beta rec
Lats gene encoding
Drosophila melanog
Drosophila melanog
Human colon cancer
cDNA encoding nove
DNA encoding novel
Disease associated
Novel protein kina
Human Ndr serine/t
Human cDNA differe
Human NDR homologu
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
DNA encoding nove
cDNA encoding nove
Human immune/haema
DNA encoding novel
Drosophila melanog
A. gossypii AG007

XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX WPI; 1996-455275/45.
 DR P-PSDB; AAW05178.
 XX
 XX New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 XX Claim 6; Page 118-123; 215pp; English.
 XX
 CC This sequence encodes a human large tumour suppressor h-lats protein,
 CC and has been isolated from a fetal human brain phage lambda-gt10
 CC cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene
 CC (AAT42119). The full-length cDNA is present in plasmid pBS(KS)-h-lats.
 CC The gene encodes a putative protein-serine/threonine-kinase, and
 CC inhibits cell proliferation and plays a crucial role throughout
 CC development. Activators or inhibitors of lats function (e.g. an
 CC antisense oligonucleotide or dominant-negative lats fragment) may be
 CC used in therapy of cancer or other proliferative disorders,
 CC degenerative disorders, trauma, growth deficiency, etc., and
 CC fragments of the gene may be used as diagnostic probes. A
 CC lats-inhibitor sequence may be expressed in a transgenic plant or
 CC farm animal to confer increased growth and inhibit senescence.
 XX
 SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;
 Query Match 100.0%; Score 3984; DB 17; Length 3984;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCTTTGGTGTGGAGGACTCTGGCCGCTCAGCGTCCGCGCTCAGCGCCGCTGGCC 60
 DB 1 ACCTTTGGTGTGGAGGACTCTGGCCGCTCAGCGTCCGCGCTCAGCGCCGCTGGCC 60
 QY 61 GCTCTCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTAAGAAATTTTA 120
 DB 61 GCTCTCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTAAGAAATTTTA 120
 QY 121 ACAGTCTCTGGGACTCTCTGAAGGATCATTTTTCACATTTTCTCAGAGAAAGCTCTGGA 180
 DB 121 ACAGTCTCTGGGACTCTCTGAAGGATCATTTTTCACATTTTCTCAGAGAAAGCTCTGGA 180
 QY 181 TCTATCAATAAAGAGTCTCTGTGGGTACATATATAGATGTTTTCATGAGAGGA 240
 DB 181 TCTATCAATAAAGAGTCTCTGTGGGTACATATATAGATGTTTTCATGAGAGGA 240
 QY 241 GTGAAAGCCAGAGGATATAGACAATGAGGCTTAAGACCTTTCTCCAGTAACATA 300
 DB 241 GTGAAAGCCAGAGGATATAGACAATGAGGCTTAAGACCTTTCTCCAGTAACATA 300
 QY 301 CTGTCACTAGCGGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 DB 301 CTGTCACTAGCGGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 QY 361 CATCTGATGCTGCTAAGCTGAGCATACATGAGTAAATGTAACCCGAGATCCTCGAC 420
 DB 361 CATCTGATGCTGCTAAGCTGAGCATACATGAGTAAATGTAACCCGAGATCCTCGAC 420
 QY 421 AAGTCAGAAATCCACCAATTTGGGCGCATATAGAGCTTTCAGGAAATTCGAAACT 480
 DB 421 AAGTCAGAAATCCACCAATTTGGGCGCATATAGAGCTTTCAGGAAATTCGAAACT 480
 QY 481 CTCTGCTTCCATTTGCAATGAAACAATTTCTCTCGAGTACATCAGAGTAAATCCAC 540
 DB 481 CTCTGCTTCCATTTGCAATGAAACAATTTCTCTCGAGTACATCAGAGTAAATCCAC 540
 QY 541 AAATGCTTCAAGACTTGAAGCTGCTGGATTTGATGAGGATATGTTTATACAGCTCTTC 600
 DB 541 AAATGCTTCAAGACTTGAAGCTGCTGGATTTGATGAGGATATGTTTATACAGCTCTTC 600
 QY 601 AGAAACTTAACACAGAGTATAGACCAATTTGAATTCATTAGTAAATGAGTTACC 660
 DB 601 AGAAACTTAACACAGAGTATAGACCAATTTGAATTCATTAGTAAATGAGTTACC 660

DB 601 AGAAACTTAACACAGAGTATAGACCAATTTGAATTCATTAGTAAATGAGTTACC 660
 QY 661 AAGATCTCTCAGCAGAGCAGATGGCTGCAGAGCTGCAGAGCTTATTATGCGCAGCATGA 720
 DB 661 AAGATCTCTCAGCAGAGCAGATGGCTGCAGAGCTGCAGAGCTTATTATGCCAGCATGA 720
 QY 721 AACCAGGGAATGTGACGAATCAGTTAACCGCAACAGAGCTGGAAAGGTTCTAAAGAT 780
 DB 721 AACCAGGGAATGTGACGAATCAGTTAACCGCAACAGAGCTGGAAAGGTTCTAAAGAT 780
 QY 781 CCTTAGTTCCTCAGAGCATGCCCGCCACTAGGAGAAAGTCTGCCCTATCATCTCTGAGA 840
 DB 781 CCTTAGTTCCTCAGAGCATGCCCGCCACTAGGAGAAAGTCTGCCCTATCATCTCTGAGA 840
 QY 841 GTCCCAACTCACAGACAGATGTAGGAACCTTTTCTGTGGATCTGTATATACAGCATTTG 900
 DB 841 GTCCCAACTCACAGACAGATGTAGGAACCTTTTCTGTGGATCTGTATATACAGCATTTG 900
 QY 901 TTCAAGCTCACCTTAGCAACGAGAGAGTGAACCCGCCACCACCTCAAGTAAGA 960
 DB 901 TTCAAGCTCACCTTAGCAACGAGAGAGTGAACCCGCCACCACCTCAAGTAAGA 960
 QY 961 GTGTACTCTCCACCACCTCCAGAGGCGGAGTCCCTCCCAAGAGGTACAACCTCCAC 1020
 DB 961 GTGTACTCTCCACCACCTCCAGAGGCGGAGTCCCTCCCAAGAGGTACAACCTCCAC 1020
 QY 1021 CTCCCTCTCATGGAGCAAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACG 1080
 DB 1021 CTCCCTCTCATGGAGCAAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACG 1080
 QY 1081 TAATCTCCGAATCTCTCTGCCACCTGGGGGAGTGAAGAGGCTATCTCTCCACCAC 1140
 DB 1081 TAATCTCCGAATCTCTCTGCCACCTGGGGGAGTGAAGAGGCTATCTCTCCACCAC 1140
 QY 1141 CTCTCAACACTTCCCGCATGATCTCTCTCAATCAAGGACAGAGAGGCTATGTTCTGTTTC 1200
 DB 1141 CTCTCAACACTTCCCGCATGATCTCTCTCAATCAAGGACAGAGAGGCTATGTTCTGTTTC 1200
 QY 1201 CTGTTGGCAGACAACCAATCATGAGAGTCTAGCAAAATTTAACTTTCCATCAGGGA 1260
 DB 1201 CTGTTGGCAGACAACCAATCATGAGAGTCTAGCAAAATTTAACTTTCCATCAGGGA 1260
 QY 1261 GACCTGGAATCGAATGCTGACAGAGTCTAGCAAAATTTAACTTTCCATCAGGGA 1320
 DB 1261 GACCTGGAATCGAATGCTGACAGAGTCTAGCAAAATTTAACTTTCCATCAGGGA 1320
 QY 1321 CTGCTGGCAGTGAATCGGAGCCACCCTCCATATCTCTGACAGCAGCTAATGAC 1380
 DB 1321 CTGCTGGCAGTGAATCGGAGCCACCCTCCATATCTCTGACAGCAGCTAATGAC 1380
 QY 1381 AAAGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTCTGCTATATACAAATGGA 1440
 DB 1381 AAAGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTCTGCTATATACAAATGGA 1440
 QY 1441 GTATTCTCAGTCTATGATGTGCAACAGAAATAGTCAATACATGGAACATATATACA 1500
 DB 1441 GTATTCTCAGTCTATGATGTGCAACAGAAATAGTCAATACATGGAACATATATACA 1500
 QY 1501 TTAGTGTACCTGGAGTGCACAAATTTGGCTCAGTCAATCTCTGCTCCAGCCAGTCA 1560
 DB 1501 TTAGTGTACCTGGAGTGCACAAATTTGGCTCAGTCAATCTCTGCTCCAGCCAGTCA 1560
 QY 1561 CCCGAGCAGTGGGATGAAATCCCTACATGGCAACCTTAACATACAGTGGGTCAAAT 1620
 DB 1561 CCCGAGCAGTGGGATGAAATCCCTACATGGCAACCTTAACATACAGTGGGTCAAAT 1620
 QY 1621 CTTTAAATACCCATAGGAAATAGCAAGTCACTCTGCTAAATTTCTCAGCCCTCTGCTA 1680
 DB 1621 CTTTAAATACCCATAGGAAATAGCAAGTCACTCTGCTAAATTTCTCAGCCCTCTGCTA 1680
 QY 1681 CAACAGTCACTGCAATTAACACAGTCTCTTATTAACAGCCTGTGAAAGTATGCTGTAT 1740
 DB 1681 CAACAGTCACTGCAATTAACACAGTCTCTTATTAACAGCCTGTGAAAGTATGCTGTAT 1740

QY 1741 TAAACCAGAGCTACAGCTGCTTTTACACCTTACACACCTTCTTGGATACACAGCCAA 1800
 DB 1741 TAAACCAGAGCTACAGCTGCTTTTACACCTTACACACCTTCTTGGATACACAGCCAA 1800
 QY 1801 TTCAAACTGTTCAACCCAGCTGCTTTTCCAGAGGAAACCGCTTCAAAATGAGCTGTATGC 1860
 DB 1801 TTCAAACTGTTCAACCCAGCTGCTTTTCCAGAGGAAACCGCTTCAAAATGAGCTGTATGC 1860
 QY 1861 CACCTGCTGCTGAAGCTTCAAACTATCAAGGACACACACCTACCCCAAAACATCTGC 1920
 DB 1861 CACCTGCTGCTGAAGCTTCAAACTATCAAGGACACACACCTACCCCAAAACATCTGC 1920
 QY 1921 TGCACCAAAACCCATCTGTTCTCCATAGCAGTCAATCAGTAAGCCTAGCAAGAGGATC 1980
 DB 1921 TGCACCAAAACCCATCTGTTCTCCATAGCAGTCAATCAGTAAGCCTAGCAAGAGGATC 1980
 QY 1981 AGCCAAAGCTTGCCCAAGAGGATGAGAGTGAAAAGATTATGAAAATGTTGATAGTGGG 2040
 DB 1981 AGCCAAAGCTTGCCCAAGAGGATGAGAGTGAAAAGATTATGAAAATGTTGATAGTGGG 2040
 QY 2041 ATAAAGAAAGAAACAGATTACAACCTCACCTTACTTACTTCTTCTCAAGCATTTAAATTTCTTTATGG 2100
 DB 2041 ATAAAGAAAGAAACAGATTACAACCTCACCTTACTTACTTCTTCTCAAGCATTTAAATTTCTTTATGG 2100
 QY 2101 AAGAGCAAGGAAATCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATTTCTTTATGG 2160
 DB 2101 AAGAGCAAGGAAATCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATTTCTTTATGG 2160
 QY 2161 AGCAACATGTAGAAAATGTACTCAATCTCATCAGCAGGCTTACATCTGTAAGAAACAT 2220
 DB 2161 AGCAACATGTAGAAAATGTACTCAATCTCATCAGCAGGCTTACATCTGTAAGAAACAT 2220
 QY 2221 TAGAGAAATGATGCGGTTGGATTATCTCAAGATCCCGAGGATCAATGAGAAAGA 2280
 DB 2221 TAGAGAAATGATGCGGTTGGATTATCTCAAGATCCCGAGGATCAATGAGAAAGA 2280
 QY 2281 TGCCTTGGCAAAAGAAATCTAATTAATACATCCGCTTTAAAGGGCTAAATGGCAAGCTTA 2340
 DB 2281 TGCCTTGGCAAAAGAAATCTAATTAATACATCCGCTTTAAAGGGCTAAATGGCAAGCTTA 2340
 QY 2341 TGTGTTGAGATTAAGACACTAGGATAGGAGCATTTGGTGAAGTCTCTAGCAAGAA 2400
 DB 2341 TGTGTTGAGATTAAGACACTAGGATAGGAGCATTTGGTGAAGTCTCTAGCAAGAA 2400
 QY 2401 AAGTAGACTAAGGCTTTGTTATGCAACAAACCTCTCGAAAGAAAGATGTTCTTCTTC 2460
 DB 2401 AAGTAGACTAAGGCTTTGTTATGCAACAAACCTCTCGAAAGAAAGATGTTCTTCTTC 2460
 QY 2461 GAAATCAAGTCTCATGTTAAGGCTGAGAGATATCTTGGCTGAAGCTGACAATGAAT 2520
 DB 2461 GAAATCAAGTCTCATGTTAAGGCTGAGAGATATCTTGGCTGAAGCTGACAATGAAT 2520
 QY 2521 GGGTAGTCTGCTATATTAATTCATCCAGATAAGGACAATTTATTAATTTGTAATGGACT 2580
 DB 2521 GGGTAGTCTGCTATATTAATTCATCCAGATAAGGACAATTTATTAATTTGTAATGGACT 2580
 QY 2581 ACATTCTGGGGTGATATGATGAGCTTAAATAGAGTGGCATCTTCCAGAAAGTC 2640
 DB 2581 ACATTCTGGGGTGATATGATGAGCTTAAATAGAGTGGCATCTTCCAGAAAGTC 2640
 QY 2641 TGGCAGATTCTACATAGCAGAACTTACCTGTGCAAGTGGAAAGTGTTCATAAAATGGGTT 2700
 DB 2641 TGGCAGATTCTACATAGCAGAACTTACCTGTGCAAGTGGAAAGTGTTCATAAAATGGGTT 2700
 QY 2701 TTATTCATAGAGATATAAACCTGATAATATTTTGAATGATCGTAGGCTCATATTAAT 2760
 DB 2701 TTATTCATAGAGATATAAACCTGATAATATTTTGAATGATCGTAGGCTCATATTAAT 2760
 QY 2761 TCAGTACCTTGGCCTCTGCTGCTTCCAGATGGACACACCTTCTAGTACTATCAGA 2820
 DB 2761 TCAGTACCTTGGCCTCTGCTGCTTCCAGATGGACACACCTTCTAGTACTATCAGA 2820

QY 2821 GTGGTGACCATCCACGGCAAGATAGCATGATTCAGTAAATGAGGGGATCCCTCAA 2880
 DB 2821 GTGGTGACCATCCACGGCAAGATAGCATGATTCAGTAAATGAGGGGATCCCTCAA 2880
 QY 2881 GCTGTCGATGTGGAGACAGACTGAAGCCATTTAGAGGGGAGAGCTGCACGCCAGCAGC 2940
 DB 2881 GCTGTCGATGTGGAGACAGACTGAAGCCATTTAGAGGGGAGAGCTGCACGCCAGCAGC 2940
 QY 2941 GATGTCATACACATCTCTTGGTGGGACTCCCAATTTATTTGACCTGAAGTCTTCTGCTAC 3000
 DB 2941 GATGTCATACACATCTCTTGGTGGGACTCCCAATTTATTTGACCTGAAGTCTTCTGCTAC 3000
 QY 3001 GAACAGATACACACAGTGTGCTGATTGGTGGAGTGTGGTGTATTCTTTTGAATGT 3060
 DB 3001 GAACAGATACACACAGTGTGCTGATTGGTGGAGTGTGGTGTATTCTTTTGAATGT 3060
 QY 3061 TGGTGGGCAACCTCTTCTTGGGCAACACACCATTTAGAAACACAAATGAAGGTTATCA 3120
 DB 3061 TGGTGGGCAACCTCTTCTTGGGCAACACACCATTTAGAAACACAAATGAAGGTTATCA 3120
 QY 3121 ACTGGCAACATCTCTTCAATTTCCACCAAGCTAACTCAGTCTGAGCTTCTGATC 3180
 DB 3121 ACTGGCAACATCTCTTCAATTTCCACCAAGCTAACTCAGTCTGAGCTTCTGATC 3180
 QY 3181 TTATTATTAACCTTTCCGAGGACCCGAAAGATCGCTTAGCAAGAATGGTCTGATGAA 3240
 DB 3181 TTATTATTAACCTTTCCGAGGACCCGAAAGATCGCTTAGCAAGAATGGTCTGATGAA 3240
 QY 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTCCAGTACCTTGAGACAGCAGCTG 3300
 DB 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTCCAGTACCTTGAGACAGCAGCTG 3300
 QY 3301 CTTTATACATCTCTTAAATCACAACCCACAGATACATCAATTTTGTCTGTTGATC 3360
 DB 3301 CTTTATACATCTCTTAAATCACAACCCACAGATACATCAATTTTGTCTGTTGATC 3360
 QY 3361 CTGATAAATTTAGGAGTGTATGATTAACGAGGAGAAATGTAATGACACTCTCAATGGAT 3420
 DB 3361 CTGATAAATTTAGGAGTGTATGATTAACGAGGAGAAATGTAATGACACTCTCAATGGAT 3420
 QY 3421 GGTATAAATTTGGAAGCATCTGACATGCAATCTTATGAATTTACCTTCCGAGGTTTT 3480
 DB 3421 GGTATAAATTTGGAAGCATCTGACATGCAATCTTATGAATTTACCTTCCGAGGTTTT 3480
 QY 3481 TTGATCACAATGGCTACCCATATAATTATCCGAAGCTTATGAATATGAATACATTAAT 3540
 DB 3481 TTGATCACAATGGCTACCCATATAATTATCCGAAGCTTATGAATATGAATACATTAAT 3540
 QY 3541 CACAAGGCTCAGACGACGCTGAGATGAAGATGATCAAAACACAGCTCAGAGATTTAAA 3600
 DB 3541 CACAAGGCTCAGACGACGCTGAGATGAAGATGATCAAAACACAGCTCAGAGATTTAAA 3600
 QY 3601 ATCGGATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAAG 3660
 DB 3601 ATCGGATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAAG 3660
 QY 3661 GGCCTCAAAATGGAGGTTTGGAGTCTGAGAGTAAATTTATGCAATATGACAGAGC 3720
 DB 3661 GGCCTCAAAATGGAGGTTTGGAGTCTGAGAGTAAATTTATGCAATATGACAGAGC 3720
 QY 3721 TATATATGCTGCTGTACAAATTTTATTTCTTAAATTTATGGAAATCTCTTTAA 3780
 DB 3721 TATATATGCTGCTGTACAAATTTTATTTCTTAAATTTATGGAAATCTCTTTAA 3780
 QY 3781 AATGTTAATTTATCCAGCGGTTTAAATCAGTATTTAGAAAAAAATTTGTTAAGGAAG 3840
 DB 3781 AATGTTAATTTATCCAGCGGTTTAAATCAGTATTTAGAAAAAAATTTGTTAAGGAAG 3840
 QY 3841 TAAATTTCAACTGAATATTTATAGTCAGTCTTGGTACTTAAAGTACTTAAATAGTAG 3900
 DB 3841 TAAATTTCAACTGAATATTTATAGTCAGTCTTGGTACTTAAAGTACTTAAATAGTAG 3900
 QY 3901 TCGTTTGTAAAAAGGAGAAACCTGCTATCTATTTCTATATATGCTATAATTTTAAAA 3960

Db 3901 TGCCTTTGTTAAAGGAGAACCTGGTATCTATTGTTATATATGCTAAATAATTTAAAA 3960
 QY 3961 TACAAGAGTTTTTGAATTTTTT 3984
 Db 3961 TACAAGAGTTTTTGAATTTTTT 3984

RESULT 2

AAZ51505
 ID AAZ51505 standard; DNA: 3984 BP.

AC AAZ51505;
 XX AAZ51505;

DT 21-JUN-2000 (first entry)

DE Human Lats (large tumour suppressor) DNA.

KW Human; Lats; large tumour suppressor; cytostatic; vulnery;
 KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
 KW treatment; prevention; screening; cancer; skin; ovarian tumour;
 KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
 KW LH; lutinizing hormone hypogonadotropic hypogonadism; metaplasia;
 KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
 KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 231..3623
 CDS /*tag= a

FT /product= "Lats protein"
 FT

PN WO200010602-A1.

PD 02-MAR-2000.

XX 18-AUG-1999; 99WO-US19068.
 XX

PR 18-AUG-1998; 98US-0096996.
 PR 18-AUG-1998; 98US-0096997.

XX (UYVA) UNIV YALE.
 PA

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
 PI Turenchalk GS, Stewart RA;

XX WPI; 2000-246496/21.
 DR P-PSDB; AAY70390.

XX Use of lats proteins, complexes of lats and cdc2 for treating cancer
 PT that is refractory to treatment by standard chemotherapy and radiation
 PT therapy, and disorders associated with aberrant levels of cdc2 activity

XX Claim 44; Fig 12; 134pp; English.

XX The present sequence is a DNA encoding human Lats (large tumour
 CC suppressor) protein which is a cell overproliferation inhibitor and a
 CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 CC The present sequence is useful for treating cancer that is refractory
 CC to standard chemotherapy or radiation therapy such as hyperplasia,
 CC metaplasia, or dysplasia, and disorders associated with aberrant
 CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
 CC include degenerative disorders, growth deficiencies, hypoproliferative
 CC disorders, physical trauma, lesions, and wounds. An animal model
 CC preferably a mouse, in which a lats gene has been disrupted by homologous
 CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. lutinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;
 Query Match 100.0%; Score 3984; DB 21; Length 3984;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTTTGGTGTCTGGGACGGACTCTGGCGGCTCAGCGCTCGCCCTCAGGCCCTGGCC 60
 Db 1 ACCTTTGGTGTCTGGGACGGACTCTGGCGGCTCAGCGCTCGCCCTCAGGCCCTGGCC 60
 QY 61 GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATATTCTAAGAAATTTA 120
 Db 61 GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATATTCTAAGAAATTTA 120
 QY 121 ACAGTCTCTGGGACTTCTTGAAGGATCATTTTTCATTTTCTCAGAAAGAGTCTGGA 180
 Db 121 ACAGTCTCTGGGACTTCTTGAAGGATCATTTTTCATTTTCTCAGAAAGAGTCTGGA 180
 QY 181 TCTATCAATAAAGAGTCCCTTCGTGTGGCTACATATATAGATGTTTTCATGAAGAGGA 240
 Db 181 TCTATCAATAAAGAGTCCCTTCGTGTGGCTACATATATAGATGTTTTCATGAAGAGGA 240
 QY 241 GTGAAAGCCAGAGGATATACAAATAGAGCCCTTTCCTGCCAGTAACATATA 300
 Db 241 GTGAAAGCCAGAGGATATACAAATAGAGCCCTTTCCTGCCAGTAACATATA 300
 QY 301 CTGTCAAGTACGGCGCAATGTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 Db 301 CTGTCAAGTACGGCGCAATGTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 QY 361 CATCTGATGCTCTAAGGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC 420
 Db 361 CATCTGATGCTCTAAGGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC 420
 QY 421 AAGTCAGAAATCCACCAATTTGGGACGATCATATAAGCCTTTCAGAGAAATTCGAAACT 480
 Db 421 AAGTCAGAAATCCACCAATTTGGGACGATCATATAAGCCTTTCAGAGAAATTCGAAACT 480
 QY 481 CTCTGCTTCCATTGCAAAATCAACAAATTTCTCGGAGTACTTTCAGAGATTAATCCAC 540
 Db 481 CTCTGCTTCCATTGCAAAATCAACAAATTTCTCGGAGTACTTTCAGAGATTAATCCAC 540
 QY 541 AAATGCTTCAAGACTTCAAGCTGCTGGATTGATGAGGATATGGTTATACAAGCTCTTC 600
 Db 541 AAATGCTTCAAGACTTCAAGCTGCTGGATTGATGAGGATATGGTTATACAAGCTCTTC 600
 QY 601 AGAAACTAACACAGAGATATAGAACGACCAATTAATTCATTAGTAAATCAGTTACC 660
 Db 601 AGAAACTAACACAGAGATATAGAACGACCAATTAATTCATTAGTAAATCAGTTACC 660
 QY 661 AAGATCTTCGACGAGACGATGGCTGCAGACGCTGCAGACCTATTAAATGCCAGCATGA 720
 Db 661 AAGATCTTCGACGAGACGATGGCTGCAGACGCTGCAGACCTATTAAATGCCAGCATGA 720
 QY 721 AACGAGGAATGTGACAGCAATCAGTTAAACGCAACAGAGCTGGAAGGTTCTTAAAGAAT 780
 Db 721 AACGAGGAATGTGACAGCAATCAGTTAAACGCAACAGAGCTGGAAGGTTCTTAAAGAAT 780
 QY 781 CCTTAGTCTCTCAGAGCATGGCCCGGCTAGGAGAAAGTGTGGCTATCATCTGAGA 840
 Db 781 CCTTAGTCTCTCAGAGCATGGCCCGGCTAGGAGAAAGTGTGGCTATCATCTGAGA 840
 QY 841 GTCCCAACTCACAGACAGATGTAGGAAGACCTTTGTCTGGATCTGTATATCAGCATTTG 900
 Db 841 GTCCCAACTCACAGACAGATGTAGGAAGACCTTTGTCTGGATCTGTATATCAGCATTTG 900
 QY 901 TTCAAGCTCACCTTAGCAGGACGAGAGTAGTGAACCCGCCACACCTCAAGTAAGA 960
 Db 901 TTCAAGCTCACCTTAGCAGGACGAGAGTAGTGAACCCGCCACACCTCAAGTAAGA 960
 QY 961 GTGTTACTCTCCACCCTTCAAGAGGCGACACTCCCTCCCAAGAGGTACAACCTCCAC 1020
 Db 961 GTGTTACTCTCCACCCTTCAAGAGGCGACACTCCCTCCCAAGAGGTACAACCTCCAC 1020

Qy	1021	CTCCCCCTTCATGGGAACCAACTCTCAAACAAGCGGTATTCTGGAACAACATGGAATACG	1080
I			
Db	1021	CTCCCCCTTCATGGGAACCAAACTCTCAAACAAGCGGTATTCTGGAACAACATGGAATACG	1080
I			
Qy	1081	TAAATCCTCCGAATCTCTCTGTGCCACCTGGGGCATGCCAAGAGGGCTATCTCTCCACCAC	1140
I			
Db	1081	TAAATCCTCCGAATCTCTCTGTGCCACCTGGGGCATGCCAAGAGGGGTATCTCTCCACCAC	1140
I			
Qy	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGAGGCATTAGTTCTGTTCT	1200
I			
Db	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGAGGCATTAGTTCTGTTCT	1200
I			
Qy	1201	CTGTTGGCAGACAACCAATCATCTGCAGAGTTCTAGCAAAATTTAACTTTCCATCAGGGA	1260
I			
Db	1201	CTGTTGGCAGACAACCAATCATCTGCAGAGTTCTAGCAAAATTTAACTTTCCATCAGGGA	1260
I			
Qy	1261	GACCTGGAATGCAAGATGGTACTGGACAACACTGATTTTCATGATACACCAAAATGTTCTGC	1320
I			
Db	1261	GACCTGGAATGCAAGATGGTACTGGACAACACTGATTTTCATGATACCAAAATGTTCTGC	1320
I			
Qy	1321	CTGCTGGCACTGTGAATCGGCAGCCACCACCTCCCATATCTCTGCACAGCAGCTAATGGAC	1380
I			
Db	1321	CTGCTGGCACTGTGAATCGGCAGCCACCACCTCCCATATCTCTGCACAGCAGCTAATGGAC	1380
I			
Qy	1381	AAGGCCCTTCTGCTTTTACAAAACAGGGGNATCTGCTGCTCTTGTCATATACAAATGGAA	1440
I			
Db	1381	AAGGCCCTTCTGCTTTTACAAAACAGGGGNATCTGCTGCTCTTGTCATATACAAATGGAA	1440
I			
Qy	1441	GTATTCTCAGTCTATGATGGTGCCAACAGAAAAATAGTCATAACATGGAACTATATAACA	1500
I			
Db	1441	GTATTCTCAGTCTATGATGGTGCCAACAGAAAAATAGTCATAACATGGAACTATATAACA	1500
I			
Qy	1501	TTAGTGTACCTGGACTGCAAAACAAATGGCCCTCAGTCTATCTTCTGCTCCAGCCCAGTCAT	1560
I			
Db	1501	TTAGTGTACCTGGACTGCAAAACAAATGGCCCTCAGTCTATCTTCTGCTCCAGCCCAGTCAT	1560
I			
Qy	1561	CCCGAGCAGTGGGCATGAATCCCTACATGCCAACCTTAAACATACCAGTGAGTCAAAATT	1620
I			
Db	1561	CCCGAGCAGTGGGCATGAATCCCTACATGCCAACCTTAAACATACCAGTGAGTCAAAATT	1620
I			
Qy	1621	CTTTTAAATRAACCCATTAGGAAATAGAGAAGTCACTCTGCTAAATCTCAGCCTTCTGCTA	1680
I			
Db	1621	CTTTTAAATRAACCCATTAGGAAATAGAGAAGTCACTCTGCTAAATCTCAGCCTTCTGCTA	1680
I			
Qy	1681	CAACAGTCTACGAAATACACAGCTCTTATTTCAACAGCCTGTGAAAGTAGCGGTGAT	1740
I			
Db	1681	CAACAGTCTACGAAATACACAGCTCTTATTTCAACAGCCTGTGAAAGTAGCGGTGAT	1740
I			
Qy	1741	TAAACACAGAGCTACAGACTGCTTTAGCACCTTACACACCCCTTCTTGATACCCACAGCAA	1800
I			
Db	1741	TAAACACAGAGCTACAGACTGCTTTAGCACCTTACACACCCCTTCTTGATACCCACAGCAA	1800
I			
Qy	1801	TTCAAACTGTTC AACCCAGTCTCTTTTCTGAGGGAACCGCTTCAAAATGTCAGCTGTGATGC	1860
I			
Db	1801	TTCAAACTGTTC AACCCAGTCTCTTTTCTGAGGGAACCGCTTCAAAATGTCAGCTGTGATGC	1860
I			
Qy	1861	CACCTGTTGCTGAAGCTCCAAACTATCAAGGACACACACACCTACCCAAAACATCTGC	1920
I			
Db	1861	CACCTGTTGCTGAAGCTCCAAACTATCAAGGACACACACACCTACCCAAAACATCTGC	1920
I			
Qy	1921	TGCACCAAAACCCATCTGTTCTCTCCATACGAGTCAATCAGTAAAGCTAGCAAGAGGATC	1980
I			
Db	1921	TGCACCAAAACCCATCTGTTCTCTCCATACGAGTCAATCAGTAAAGCTAGCAAGAGGATC	1980
I			
Qy	1981	AGCCAAGCTTGGCCCAAGGAAGTAGAGTGAAGAGAGTTATGAAAATGTTGATAGTGGGG	2040
I			
Db	1981	AGCCAAGCTTGGCCCAAGGAAGTAGAGTGAAGAGAGTTATGAAAATGTTGATAGTGGGG	2040
I			
Qy	2041	ATAAGAAAGAAGAACAGATTACAACTTCACCTATTACTGTTAGGAAAACAAAGAGATG	2100
I			
Db	2041	ATAAGAAAGAAGAACAGATTACAACTTCACCTATTACTGTTAGGAAAACAAAGAGATG	2100
I			

QY	2101	AGAGCGAAGGAACTCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG	2160
DB	2101	AAAGCGAAGGAACTCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG	2160
QY	2161	AGCAACATCTAGAAAATGTACTCAAATCTCATCAGCAGCGTTACATCGTAAAAACAAT	2220
DB	2161	AGCAACATCTAGAAAATGTACTCAAATCTCATCAGCAGCGTTACATCGTAAAAACAAT	2220
QY	2221	TAGAGAAATCAAAATGATCGCGGTTGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGA	2280
DB	2221	TAGAGAAATCAAAATGATCGCGGTTGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGA	2280
QY	2281	TGCTTTGCCAAAAGAAATCTAAATACATCCGTCCTTAAAAAGGCTAAAAATGGACAAGCTCTA	2340
DB	2281	TGCTTTGCCAAAAGAAATCTAAATACATCCGTCCTTAAAAAGGCTAAAAATGGACAAGCTCTA	2340
QY	2341	TGTTTGTGAAGATAAGACACATAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAA	2400
DB	2341	TGTTTGTGAAGATAAGACACATAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAA	2400
QY	2401	AAATAGATACTAAGCGTTTGTATGCAACAAAAAATCTTCGAAAGAAAGATGTTCTTCTTC	2460
DB	2401	AAATAGATACTAAGCGTTTGTATGCAACAAAAAATCTTCGAAAGAAAGATGTTCTTCTTC	2460
QY	2461	GAATCAAGTCGCTCATGTTAAAGGCTGAGAGATATCCTGGCTGAAGCTGACAATGAAT	2520
DB	2461	GAATCAAGTCGCTCATGTTAAAGGCTGAGAGATATCCTGGCTGAAGCTGACAATGAAT	2520
QY	2521	GGGTAGTTGCTCTATATTATTTCATTCCAAGATAAGACAAATTTATACTTTGTAATGGACT	2580
DB	2521	GGGTAGTTGCTCTATATTATTTCATTCCAAGATAAGACAAATTTATACTTTGTAATGGACT	2580
QY	2581	ACATCTCTGGGGTGATATGATGAGCGCTATTAATAGAAATGGGCATCTTTCAGAAAGTC	2640
DB	2581	ACATCTCTGGGGTGATATGATGAGCGCTATTAATAGAAATGGGCATCTTTCAGAAAGTC	2640
QY	2641	TGGCAGATTTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTT	2700
DB	2641	TGGCAGATTTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTT	2700
QY	2701	TTATTATAGAGATATTAAACCTGATAATATTTTGATTGATCGTGATGTCATATTAAAT	2760
DB	2701	TTATTATAGAGATATTAAACCTGATAATATTTTGATTGATCGTGATGTCATATTAAAT	2760
QY	2761	TGACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACAGATTTCAAGTACTATCAGA	2820
DB	2761	TGACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACAGATTTCAAGTACTATCAGA	2820
QY	2821	GTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAATGAATGGGGGATCCCTCAA	2880
DB	2821	GTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAATGAATGGGGGATCCCTCAA	2880
QY	2881	GCTGTGATGTGGAGACAGACTGAAGCCATTAGCGGGAGAGCTGCACGCCAGCACCGC	2940
DB	2881	GCTGTGATGTGGAGACAGACTGAAGCCATTAGCGGGAGAGCTGCACGCCAGCACCGC	2940
QY	2941	GATGCTTAGCACATCTTTGGTTGGGACCTCCCAATTTATTCACCTGAAGTGTTCCTAC	3000
DB	2941	GATGCTTAGCACATCTTTGGTTGGGACCTCCCAATTTATTCACCTGAAGTGTTCCTAC	3000
QY	3001	GAACAGGATACACAGTGTGTGATTGGTGGAGTGTGGTGTTATCTTTTGAATGT	3060
DB	3001	GAACAGGATACACAGTGTGTGATTGGTGGAGTGTGGTGTTATCTTTTGAATGT	3060
QY	3061	TGGTGGGACAACCTCTCTTCTTTGGCACAACACCATTAGAACAACAATGAAGGTTATCA	3120
DB	3061	TGGTGGGACAACCTCTCTTCTTTGGCACAACACCATTAGAACAACAATGAAGGTTATCA	3120
QY	3121	ACTGGCAACAATCTCTTACATTCACCACACAGCTAAACTCAGTCTCTGAAGCTTCTGATC	3180
DB	3121	ACTGGCAACAATCTCTTACATTCACCACACAGCTAAACTCAGTCTCTGAAGCTTCTGATC	3180
QY	3181	TTATTATTAAACTTTGCCGAGGACCCGAAGATCGCTTAGCGAAGAAATGGTCTGATGAAA	3240

```

Db 3181 TTATTATTAACTTTCCGAGGACCGAAGATCGCTAGGCAAGAATGGTCTGATGAA 3240
Qy 3241 TAAAGCTCATCATTTTAAACAAATTCACCTCCAGTACCTCAGACACAGCTCTG 3300
Db 3241 TAAAGCTCATCATTTTAAACAAATTCACCTCCAGTACCTCAGACACAGCTCTG 3300
Qy 3301 CTTCAATACATCTCTAAATACACACACACACACACATCATCAAAATTTTTCCTGTTGATC 3360
Db 3301 CTTCAATACATCTCTAAATACACACACACACACACATCATCAAAATTTTTCCTGTTGATC 3360
Qy 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAATGACACTCTCAATGGAT 3420
Db 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAATGACACTCTCAATGGAT 3420
Qy 3421 GGTATAAATGGAAGACATCTTCAACATCATCTATGAATTTACCTCCGAAGGTTTT 3480
Db 3421 GGTATAAATGGAAGACATCTTCAACATCATCTATGAATTTACCTCCGAAGGTTTT 3480
Qy 3481 TTGATGACATGGCTACCCATATAATTTATCCGAGGCTATTGATATGATATCAATTAAT 3540
Db 3481 TTGATGACATGGCTACCCATATAATTTATCCGAGGCTATTGATATGATATCAATTAAT 3540
Qy 3541 CACAAGGCTCAGACGACGCTCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAATA 3600
Db 3541 CACAAGGCTCAGACGACGCTCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAATA 3600
Qy 3601 ATCCGATCTAGTATATGTTTAACACACTAGTAAATGAATGATGAGGATTTCTAAAG 3660
Db 3601 ATCCGATCTAGTATATGTTTAACACACTAGTAAATGAATGATGAGGATTTCTAAAG 3660
Qy 3661 GGCCTGAATGCGAGGCTGTTTGGAGTCTGAGAGTAAATTTATGCAATATGACAGAGC 3720
Db 3661 GGCCTGAATGCGAGGCTGTTTGGAGTCTGAGAGTAAATTTATGCAATATGACAGAGC 3720
Qy 3721 TATATATGTCGCTCTGTACAAATTTTATTTTCTTAATATGGAATCCTTTTAA 3780
Db 3721 TATATATGTCGCTCTGTACAAATTTTATTTTCTTAATATGGAATCCTTTTAA 3780
Qy 3781 AATGTTAATTTATCCAGCGCTTTTAAATCAGTATTTAGAAAAATTTGTTATAGGAAAG 3840
Db 3781 AATGTTAATTTATCCAGCGCTTTTAAATCAGTATTTAGAAAAATTTGTTATAGGAAAG 3840
Qy 3841 TAAATATGAATGAATATATAGTCACTTCTGCTTAAAGTAACTTAAATTAAGTAG 3900
Db 3841 TAAATATGAATGAATATATAGTCACTTCTGCTTAAAGTAACTTAAATTAAGTAG 3900
Qy 3901 TCGTTTGTAAAAAGGAGAACCTGGTATCTATTTGTATATATCTAATAATTTTAA 3960
Db 3901 TCGTTTGTAAAAAGGAGAACCTGGTATCTATTTGTATATATCTAATAATTTTAA 3960
Qy 3961 TACAAGAGTTTTTGAATTTTTTTT 3984
Db 3961 TACAAGAGTTTTTGAATTTTTTTT 3984

```

RESULT 3

AA87396

ID AA87396 standard; cDNA; 7382 BP.

AC AA87396;

DT 08-OCT-1999 (first entry)

DE Human WART1 cDNA.

KW WART1; WART1; WART orthologue; human; signal transduction;

OS Homo sapiens.

FT 5'UTR. 1...399

Key Location/Qualifiers

FT 5'UTR. 1...399

```

FT repeat_region a
FT 12..63 b
FT /*tag=
FT /*tag=
FT /note= "contains 10 copies of GGC repeat, similar
FT to repeats that undergo expansion in human
FT diseases associated with neuronal
FT phenotypes."
FT 395...3787
FT /*tag= c
FT 3788..7382
FT /*tag= d
FT 6058..6346
FT /*tag= e
FT /note= "Alu-J subfamily repeat motif"
FT 7339..7344
FT /*tag= f
FT replace(978,G)
FT /*tag= g
FT /note= "results in Ala/Gly change"
FT replace(1840)
FT /*tag= h
FT /note= "silent polymorphism"
FT replace(3252..3253)
FT /*tag= i
FT /note= "deletion of 2 adenines, causes
FT frameshift"

```

WO9937787-A2.

29-JUL-1999:

20-JAN-1999: 99WO-US01145.

21-JAN-1998: 98US-0072023.

(SUG-.) SUGEN INC.

Planagan P, Plowman GD;

WPI; 1999-458698/38.

P-PSDB; AAY06526.

New nucleic acid encoding human orthologs of Drosophila WART proteins, used to identify specific modulators for treating cancer or for diagnosis

Claim 1; Page 120-121; 137pp; English.

This is the nucleotide sequence of a cDNA clone coding for a human orthologue, i.e. WART1 (see AAY06526), of Drosophila non-receptor serine/threonine kinase WART. WART1 cDNA was isolated from a human bone marrow cDNA library using PCR fragments of WART1 as probes. 3 Polymorphisms were identified in the WART1 gene: (1) at nucleotide 978, resulting in an Ala/Gly change; (2) at nucleotide 1840, silent; and (3) at nucleotides 3252-3253, comprising a deletion of 2 adenines, resulting in a C-terminal truncation of WART1 in the putative kinase domain. The latter frameshift mutation was observed in 2 independent clones from human bone marrow cDNA. Truncation of WART1 could play a role in disease progression. WART1 shows strong expression in cell lines from non-small cell lung cancer, ovarian tumours, central nervous system tumours, renal tumours and breast tumours, and may provide a target for oncology drug development. Nucleic acids encoding full-length WART1 and WART1 polypeptides lacking one or more of amino acid segments 12-45, 55-151, 236-377, 404-520, 555-559, 601-702, 691-998 and 1011-1086, or lacking one or more of the N-terminal domain, catalytic domain, or C-terminal domain are claimed, as well as WART2 sequences (AA87397), WART1 and WART2 polypeptides, antibodies, a method for identifying modulators of WART function, and use of such modulator compounds to treat an abnormal condition involving WART signal transduction, especially cancer. Probes for detection of WART nucleic acids are also claimed.

XX

0.

Qy	1023	CCCCCTTCATGGGAACCAAACTCTCAACAAGACCGCTATTCTGGAAACATGGAAATACGTTA	1082
Db	1187	CCCCCTTCATGGGAACCAAACTCTCAACAAGACCGCTATTCTGGAAACATGGAAATACGTTA	1246
Qy	1083	ATCTCCCGAATCTCTCTGTCGCCACCTGGGGCATGGCAAGAGGGCTATCCCTCCACCACTCT	1142
Db	1247	ATCTCCCGAATCTCTCTGTCGCCACCTGGGGCATGGCAAGAGGGCTATCCCTCCACCACTCT	1306
Qy	1143	CTCAACACTTCCCCCATGAATCCTCTTAATCAAGGACAGAGAGGCAATTAGTTCGTTCCT	1202
Db	1307	CTCAACACTTCCCCCATGAATCCTCTTAATCAAGGACAGAGAGGCAATTAGTTCGTTCCT	1366
Qy	1203	GTGGCAGACAACCAATCATCATGAGAGTCTACCAATTTAACTTTCATACAGGAGA	1262
Db	1367	GTGGCAGACAACCAATCATCATGAGAGTCTACCAATTTAACTTTCATACAGGAGA	1426
Qy	1263	CCTGGAATGAGAAATGGTACTGGCAAACTGATTTTCATGATACACCAAAATGTTGTCCT	1322
Db	1427	CCTGGAATGAGAAATGGTACTGGCAAACTGATTTTCATGATACACCAAAATGTTGTCCT	1486
Qy	1323	GCTGGCACTGTGAATGGGAGCGCCACCACTCCATATCTCTGACAGACGCTAATGGACAA	1382
Db	1487	GCTGGCACTGTGAATGGGAGCGCCACCACTCCATATCTCTGACAGACGCTAATGGACAA	1546
Qy	1383	AGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTGGTTCATATACAAATGGAAGT	1442
Db	1547	AGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTGGTTCATATACAAATGGAAGT	1606
Qy	1443	ATTCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATACATGGAACATATATAACATT	1502
Db	1607	ATTCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATACATGGAACATATATAACATT	1666
Qy	1503	AGTGTACTGGACTGCAAAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCCACTCATCC	1562
Db	1667	AGTGTACTGGACTGCAAAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCCACTCATCC	1726
Qy	1563	CCGAGCAGTGGGCAATGAAATCCCTACATGGCAACCTAACTACCACTGAGGTCAAATTCCT	1622
Db	1727	CCGAGCAGTGGGCAATGAAATCCCTACATGGCAACCTAACTACCACTGAGGTCAAATTCCT	1786
Qy	1623	TTTAAATAACCCATTAGGAATAGCAAGTCACTCTGCTAATTTCTCAGCCCTTCGTCTACA	1682
Db	1787	TTTAAATAACCCATTAGGAATAGCAAGTCACTCTGCTAATTTCTCAGCCCTTCGTCTACA	1846
Qy	1683	ACAGTCACTGCAATTAACACAGCTCCTATTTCACAGCCCTGTGAAAGATATGCGGTATTA	1742
Db	1847	ACAGTCACTGCAATTAACACAGCTCCTATTTCACAGCCCTGTGAAAGATATGCGGTATTA	1906
Qy	1743	AAACAGAGCTACAGACTGCTTTAGCACCTACACACCCCTTCTTGATACACAGCCCAATT	1802
Db	1907	AAACAGAGCTACAGACTGCTTTAGCACCTACACACCCCTTCTTGATACACAGCCCAATT	1966
Qy	1803	CAAACTGTTCAACCCAGTCCCTTTTCTGAGGAAACCGCTTCAAAATGTGACTGTGATGCCA	1862
Db	1967	CAAACTGTTCAACCCAGTCCCTTTTCTGAGGAAACCGCTTCAAAATGTGACTGTGATGCCA	2026
Qy	1863	CCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCAACCCCTACCCAAACATCTCGTG	1922
Db	2027	CCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCAACCCCTACCCAAACATCTCGTG	2086
Qy	1923	CACCAAAACCCATCTGTTCTCCATCAGAGTCAATCAGTAAAGCTACCAAGAGGATCAG	1982
Db	2087	CACCAAAACCCATCTGTTCTCCATCAGAGTCAATCAGTAAAGCTACCAAGAGGATCAG	2146
Qy	1983	CCAAGCTTGCCCAAGGAAGATGAGAGTGAAGAGATTATGAAATGTTGATAGTGGGAT	2042
Db	2147	CCAAGCTTGCCCAAGGAAGATGAGAGTGAAGAGATTATGAAATGTTGATAGTGGGAT	2206
Qy	2043	AAAGAAAGAAACAGATTACAATTCACCTCACTTACTGTTAGGAAACAAACAGAAAGTAA	2102
Db	2207	AAAGAAAGAAACAGATTACAATTCACCTCACTTACTGTTAGGAAACAAACAGAAAGTAA	2266

Qy	2103	GAGCGAAGGGAATCTCGTATTCCTCAAAAGTTATTCCTCCTCAAGCATTTAAATCTCTTTATGCGAG.	2166
Db	2267	GAGCGAAGGGAATCTCGTATTCCTCAAAAGTTATTCCTCCTCAAGCATTTAAATCTCTTTATGCGAG	2326
Qy	2163	CACATGTTAGAAAATCTACTCAAAATCTCATCAGCAGCGTCTTACATCGTAAAAACAATTA	2222
Db	2327	CACATGTTAGAAAATCTACTCAAAATCTCATCAGCAGCGTCTTACATCGTAAAAACAATTA	2386
Qy	2223	GAGATGAATGATGCGGGTTCGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGATG	2282
Db	2387	GAGATGAATGATGCGGGTTCGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGATG	2446
Qy	2283	CTTTGCCAAAAAGAATCTTAATTACATCCGCTCTTAAAGGGCTAAAAATGGACAAGTCTATG	2342
Db	2447	CTTTGCCAAAAAGAATCTTAATTACATCCGCTCTTAAAGGGCTAAAAATGGACAAGTCTATG	2506
Qy	2343	TTTGTGAAGATAAGACACATAGGAATAGAGGATTTGGTGAAGTCTCTCTAGCAAGAAAA	2402
Db	2507	TTTGTGAAGATAAGACACATAGGAATAGAGGATTTGGTGAAGTCTCTCTAGCAAGAAAA	2566
Qy	2403	GTAGATACTAAGGCTTTGTGTATCAACAACAAACTCTTCGAAAGAAGATGTTCTTCTTCGA	2462
Db	2567	GTAGATACTAAGGCTTTGTGTATCAACAACAAACTCTTCGAAAGAAGATGTTCTTCTTCGA	2626
Qy	2463	AATCAAGTCGCTCATGTTAAGGCTCGAGAGATATCTCTGGCTGAAGCTGACAATGAATGG	2522
Db	2627	AATCAAGTCGCTCATGTTAAGGCTCGAGAGATATCTCTGGCTGAAGCTGACAATGAATGG	2686
Qy	2523	GTAGTTCGCTATATTAATTCATTCACAAGATGAAGACAATTTATACTTTGTAATGGACTAC	2582
Db	2687	GTAGTTCGCTATATTAATTCATTCACAAGATGAAGACAATTTATACTTTGTAATGGACTAC	2746
Qy	2583	ATTCTCTGGGGTGATATGATGAGCCTATTAATTAGAATGGCATCTTTCAGAAAGCTCG	2642
Db	2747	ATTCTCTGGGGTGATATGATGAGCCTATTAATTAGAATGGCATCTTTCAGAAAGCTCG	2806
Qy	2643	GCACGATTCATACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTT	2702
Db	2807	GCACGATTCATACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTT	2866
Qy	2703	ATTCTATAGAGATTAATAACCTCGATAATTTTGATTGATCGTGATGGTCATATTAATTTG	2762
Db	2867	ATTCTATAGAGATTAATAACCTCGATAATTTTGATTGATCGTGATGGTCATATTAATTTG	2926
Qy	2763	ACTGACTTTGGGCTCTGCACCTGGGCTCAGATGGACACACAGATTCCTAAGTACTATCAGAGT	2822
Db	2927	ACTGACTTTGGGCTCTGCACCTGGGCTCAGATGGACACACAGATTCCTAAGTACTATCAGAGT	2986
Qy	2823	GGTGACCATCCACGCCAAGATAGCATGGATTTTCAGTAATGAATGGGGGATCCCTCAAGC	2882
Db	2987	GGTGACCATCCACGCCAAGATAGCATGGATTTTCAGTAATGAATGGGGGATCCCTCAAGC	3046
Qy	2883	TCGTCCATGTGGAGACAGACTCAAGCCATTAGACGGGAGAGCTCCAGCCAGACACACGGA	2942
Db	3047	TCGTCCATGTGGAGACAGACTCAAGCCATTAGACGGGAGAGCTCCAGCCAGACACACGGA	3106
Qy	2943	TGCTATAGCACATCTTTTGTTGGGACCTCCCAATTAATTGACCTCAAGTGTGTGTACGA	3002
Db	3107	TGCTATAGCACATCTTTTGTTGGGACCTCCCAATTAATTGACCTCAAGTGTGTGTACGA	3166
Qy	3003	ACAGGATACACAGTGTGTGATTGGTGGAGTGTGGTGTATTCTTTTTGAAATGTTG	3062
Db	3167	ACAGGATACACAGTGTGTGATTGGTGGAGTGTGGTGTATTCTTTTTGAAATGTTG	3226
Qy	3063	GTGGGACAACCTCTTTCTTGGCCAAACACCATTTAGAAACACAAATGAAGGTTATCAAC	3122
Db	3227	GTGGGACAACCTCTTTCTTGGCCAAACACCATTTAGAAACACAAATGAAGGTTATCAAC	3286
Qy	3123	TGGCAAAACATCTTTCATTTCCACCAAGCTAAACTCAGTCCCTGAAGCTTCTCATCTT	3182
Db	3287	TGGCAAAACATCTTTCATTTCCACCAAGCTAAACTCAGTCCCTGAAGCTTCTCATCTT	3346
Qy	3183	ATTATTAAACTTTTCCGGAGGACCCGAAGATCGCTTTAGGCAAGAATGGTGCTGATGAAATA	3242

Db	3347	ATTATTAACCTTTCCGAGACCCGAGATCGCTTAGGCAAGAAATGGTGCTGATGAATA	3406
Qy	3243	AAAGCTCATCTTTTAAAACAATTTGACTTCTCCAGTGACCTCAGACAGCAGTCTGCT	3302
Db	3407	AAAGCTCATCTTTTAAACANATGACTTCTCCAGTGACCTCAGACAGCAGTCTGCT	3466
Qy	3303	TCATACATTCCTAAAAATCACACACCCCAACAGATACATAAAATTTTGATCCTGTTGATCCT	3362
Db	3467	TCATACATTCCTAAAAATCACACACCCCAACAGATACATAAAATTTTGATCCTGTTGATCCT	3526
Qy	3363	GATAAATTTGGAGTGAATACGAGGAGAAATGTAATGACACTCTCAATGGATGG	3422
Db	3527	GATAAATTTGGAGTGAATACGAGGAGAAATGTAATGACACTCTCAATGGATGG	3586
Qy	3423	TATAAAATGAAAGCATCTGGAACATGCTATGAATTTTACCTCCGAAAGCTTTT	3482
Db	3587	TATAAAATGAAAGCATCTGGAACATGCTATGAATTTTACCTCCGAAAGCTTTT	3646
Qy	3483	GATGACAATGGCTACCCATATAATTTATCCGAAGCCTATTGAATGAATACATAATTC	3542
Db	3647	GATGACAATGGCTACCCATATAATTTATCCGAAGCCTATTGAATGAATACATAATTC	3706
Qy	3543	CAAGGCTCAGAGCAGTCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAANAAT	3602
Db	3707	CAAGGCTCAGAGCAGTCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAANAAT	3766
Qy	3603	CGCATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGGATTTGTAAGAGG	3662
Db	3767	CGCATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGGATTTGTAAGAGG	3826
Qy	3663	CCTGAATGCGAGGTGTTTGGAGTCTCAGAGTAAAAATATGCAAAATATGACAGAGCTA	3722
Db	3827	CCTGAATGCGAGGTGTTTGGAGTCTCAGAGTAAAAATATGCAAAATATGACAGAGCTA	3886
Qy	3723	TATATGTGCTCTGTGTCACATATTTTATTTTCTCTAAATTTATGGAAATCCCTTTAAAA	3782
Db	3887	TATATGTGCTCTGTGTCACATATTTTATTTTCTCTAAATTTATGGAAATCCCTTTAAAA	3946
Qy	3783	TGTTAAATTTATCCAGCCGTTTAAATCAGTATTTAGAAAAAAATTTGTTAAAGGAAAGTA	3842
Db	3947	TGTTAAATTTATCCAGCCGTTTAAATCAGTATTTAGAAAAAAATTTGTTAAAGGAAAGTA	4006
Qy	3843	AAATATCAACTGAATATATAGTCAGTCTTCGGTACTTAAAGTACTTAAATTAAGTAGTG	3902
Db	4007	AAATATCAACTGAATATATAGTCAGTCTTCGGTACTTAAAGTACTTAAATTAAGTAGTG	4066
Qy	3903	CTTTGTTTAAAGGAGAAACCTTGGTATCTATTGTTGATATATGCTAAATAATTTTAAATA	3962
Db	4067	CTTTGTTTAAAGGAGAAACCTTGGTATCTATTGTTGATATATGCTAAATAATTTTAAATA	4126
Qy	3963	CAAGAGTTTTTGAATTTTTT	3984
Db	4127	CAAGAGTTTTTGAATTTTTT	4148
RESULT 4			
AAZ51506			
ID AAZ51506 standard; DNA; 3213 bp.			
XX			
AC	AAZ51506;		
XX			
DT	21-JUN-2000 (first entry)		
DE	Mouse Lats (large tumour suppressor) DNA.		

RESULT 4

RESOLUT 4
AAZ51506

ID AAZ51506

XX

AC AAZ51506;

XX

DT 21-JUN-2000 (first entry) VV

XX
DE
Mouse Iats (large tumour suppressor) DNA

DE mouse laccs (large tumour suppressor) DNA: XX.

KW Mouse; Lats; large tumour suppressor: cytostatic: vulnarary:

KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase

KW treatment; prevention; screening; cancer; skin; ovarian tumour;

KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia

LH; luteinizing hormone-hypogonadotropic hypogonadism; metaplasia

KW dysplasia; degenerative disorder; growth deficiency; physical trauma

KW hypoproliferative disorder: lesion: wound: late knock-out mouse: de-

kw, hypoproliferative disorder; lesion; wound; tail knock-out mouse; ds

XX OS Mus musculus.
 XX PH Key Location/Qualifiers
 XX FT CDS 1..2889
 XX FT /*tag= a
 XX FT /product= "Lats protein"
 XX FT /partial
 XX PN WO200010602-A1.
 XX PD 02-MAR-2000.
 XX PF 18-AUG-1999; 99WO-US19068.
 XX PR 18-AUG-1998; 98US-0096996.
 XX PR 18-AUG-1998; 98US-0096997.
 XX PA (UYU) UNIV YALE.
 XX XU T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
 XX PI Turenchalk GS, Stewart RA;
 XX DR WPI: 2000-246496/21.
 XX DR P-PSDB: AAY70391.
 XX PT Use of lats proteins, complexes of lats and cdc2 for treating cancer
 XX PT that is refractory to treatment by standard chemotherapy and radiation
 XX PT therapy, and disorders associated with aberrant levels of cdc2 activity.
 XX PS Claim 44; Fig 13; 134pp: English.
 XX CC The present sequence is a DNA encoding mouse Lats (large tumour
 XX CC suppressor) protein which is a cell overproliferation inhibitor and a
 XX CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 XX CC The present sequence is useful for treating cancer that is refractory
 XX CC to standard chemotherapy or radiation therapy such as hyperplasia,
 XX CC metaplasia, or dysplasia, and disorders associated with aberrant
 XX CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
 XX CC include degenerative disorders, growth deficiencies, hypoproliferative
 XX CC disorders, physical trauma, lesions, and wounds: An animal model
 XX CC preferably a mouse, in which a lats gene has been disrupted by homologous
 XX CC recombination; e.g. a lats knock-out mouse, is used for screening
 XX CC compounds that can be used to treat or prevent cancer, particularly
 XX CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 XX CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 XX CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX SQ Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;
 Query Match 62.38; Score: 2483; DB 21; Length 3213;
 Best Local Similarity 87.48; Pred. No. 0;
 Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
 Oy 732 GTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAAGTTCTTAAGAAATCCTTAGTCTCT 791
 Db 1 GTGCAACATTCATTACCGGAACAAAGCTGGAAGTTCTTAAGAGTCTCTAGTCTCT 60
 Oy 792 CAGAGGATGCGCGCCCACTAGGAGAAAGTGTGGCCTATATCTGAGAGTCCCAACTCA 851
 Db 61 CAGAGACACGGCCCATCTCTAGGAGAAAATGTGGTTTATCTGTTCTGAAAGCCCAACTCA 120
 Oy 852 CAGACAGATGAGNAGACCTTTGTCTGGATCTGGTATATCAGATTTGTTCAAGCTCAC 911
 Db 121 CAGCGGATGAGGAAGACCTCTGTCTGGATCCGGATTCAGCATTTGCTCAAGCTCAC 180
 Oy 912 CCTAGCAACGACAGAGAGTGAACCCCCACACACCTCAAGTAAGGAGTGTACTCTCT 971
 Db 181 CCAGCAATGACAGAGATGACCCCCACACCTCAAGTTAGGAGTGTACTCTCT 240
 Oy 972 CCACCACCTCCAGAGGCCAGCTCCGCCCTCCAAAGAGGTACAACCTCCACCTCCGCCCTTCA 1031
 Db 1318 AACAGATTACAACCTTCACTTATCAGTGTTCGGGAAAACAAAGAGATGAAGACGAAGA 1377

Db 241 CCACCACCTCCGAGAGGCCAGACCCACCTCCCGAGGACACCACTCCCCCTCCCTCA 300
 Oy 1032 TGGGAACCAAACTCTCAACAAAGCGCTATTTCTGAAACATCGAATAGTATCTCCCGA 1091
 Db 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTAGTATCTCCCGA 360
 Oy 1092 ATCTCTCTCTCCCACTCTGGGCGATGGCAAGAGCGCTATCTCCACCACTCTCAACACT 1151
 Db 361 ATCTCCCTCTTCCACCTCTGGGCGTGGCAGGAGGCTACCTCCACCACTCTTACCCT 420
 Oy 1152 TCCCCCATGAATCTCTCTTAATCAAGACAGAGAGCATAGTTCTGTGTTCTGTGGCAGA 1211
 Db 421 TCTCCCATGAATCTCCCTTAGCCAGCTCAGAGGCGCATAGTTCTGTCTCCAGTGTAGA 480
 Oy 1212 CAACCAATCATCATGACAGTTCTAGCAAAATTAATCTTCCATCAGGAGGACCTCGAATG 1271
 Db 481 CAACCATCATCATGACAGTACTAGCAAAATTAATCTTACACAGGCGGACCTGGAGTT 540
 Oy 1272 CAGAAATGCTACTGGACAACATGATTTATGATACACCAAAATGTTGTCTCTCTGGCACT 1331
 Db 541 CAGAAATGCTGTGTGTCAGTCTGATTTATCTGTGCACCAAAA--TGTCCCACTGTGTTCT 597
 Oy 1332 GTGAATCGGAGCCACCACTCCATATCTCTGACAGAGCTAATGAGCAAAAGCCCTTCT 1391
 Db 598 GTGACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGAGCAAAAGCCCTCT 657
 Oy 1392 GCTTTACAAACAGGGGATCTGCTCTCTCTGTCATATACAAATGGAAGTATCTCTCAG 1451
 Db 658 GCTTTACAAACAGGGGCTTCTGCTCTCCACATCATTCGCCCAATGGAAGCTTCTCTCAG 717
 Oy 1452 TCTATGATGGTCCCAACAGAAATAGTCAATACATGGAACATATAACATAGTGTACT 1511
 Db 718 TCGATGATGGTGCCCAACAGGACAGTCAATACATGGAGCTTTATATATTAATGTCCT 777
 Oy 1512 GGACTGCAACAAATGTCCTCAGTCTCTCTGTCAGGCCAGTCAATCCCGAGCAGT 1571
 Db 778 GGACTGCAACAGCGCTGGCCCGCAGTCTCTCTGCTCTGCGAGTCAATCCCAAGCGT 837
 Oy 1572 GGGCATGAATCCCTACATGCACACCTAACATACCACTGAGTCAAAATCTTTAATAAC 1631
 Db 838 GGGCATGAATTCCTACATGGCAACCTAACATACCACTGAGGTCAAAATCTTTAATAAC 897
 Oy 1632 CCATTAGGAAATAGAGCAAGTCACTCTCTCTAATCTCAGCCTTCTGCTACACAGTCACT 1691
 Db 898 CCATTAGGAAATAGAGCAAGTCACTCTCTAATCTCAGCCTTCTGCTACAGTCACT 957
 Oy 1692 GCAATTACACAGCTCTCTTATCAACAGCTGTGAAAAGTATGCGGTGTAATTAACACAG 1751
 Db 958 GCCATCACACCGCTCTCTTATCAACAGCCGCTGAAAAGCATGCGCGTCTGAAAACACAG 1017
 Oy 1752 CTACAGACTGCTTTAGCACCTTACACACCTTCTTGGATACACAGCCAAATTCACACTGTT 1811
 Db 1018 CTGACAGTGTCTTAGCCCCAACCCATCTCTTGGATGCCCACAGCCAGTTCAGACTGTT 1077
 Oy 1812 CAACCCAGTCTCTTCTCTGAGGAAACCCCTTCAATATGACTGTGATGCCACCTGTTGCT 1871
 Db 1078 CAGCTTACCCCTTTCTCTGAGGTTACAGCTTCAAGTGTGCTCTGCTCCACCTGTTGCT 1137
 Oy 1872 GAAGTCCAAACTATCAAGGACCACTCCACCTTACCCAAACATCTGCTGACCCAAAC 1931
 Db 1138 GAAGTCCAACTATCAAGTCCACCTCCACCTTATCCAAACATCTGCTACACCAAAAC 1197
 Oy 1932 CCATCTGTTCTCTCAGTCAATCAGTAAGCTTACCAAGAGGATCAGCCCAAGCTTG 1991
 Db 1198 CCATCTGTTCTCTCAGTCAATCAGTAAGCTTACCAAGAGGATGAAACAGCTTACT 1257
 Oy 1992 CCCAAGGAAGTATGAGTGAAGAGTATGAAAATGTTGATGTTGGGATTAAGAAAAG 2051
 Db 1258 CCCAAGGAGATGATGAGAGAGGTCGCGACAGTGTGACTCTGGGATTAAGAAAAG 1317
 Oy 2052 AAACAGATTACAACCTTCACTTATCTGTTAGGAAAACAAAGAGATGAAGCGAAG 2111
 Db 1318 AAACAGATTACAACCTTCACTTATCAGTGTTCGGGAAAACAAAGAGATGAAGACGAAGA 1377

[illegible][illegible]

Db 1558 AAAGAGCTAATATATTCGCTCTTAAAAAGGCTAAAAATGGACAGCTATATGTTGTAAG 1617
 Qy 2352 ATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTCTAGCAAGAAAAAGTAGATACT 2411
 Db 1618 ATAAAGACATAGGAATAGGAGCGTTTGGTGAAGTCTCTAGCAAGAAAAAGTAGATACT 1677
 Qy 2412 AAGGCTTTGATGCAACAAAAACCTCTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTC 2471
 Db 1678 AAGGCTTTGATGCAACAAAAACCTCTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTC 1737
 Qy 2472 GCTCATGTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACATGAATGGTAGTTCGT 2531
 Db 1738 GCTCATGTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACATGAATGGTAGTTCGT 1797
 Qy 2532 CTATATTTATTCATTCACAGATTAAGACAAATTTATATCTTGTGAATGGACTACATCTCTGGG 2591
 Db 1798 CTGTAATCTCTTCCAGGACAGGACAACTGTGACTTTGTGATGGACTACATCTCTGGG 1857
 Qy 2592 GGTGATATGATGAGCCTATTAATTAAGATGGCATCTTTCAGAAAGTCTGGCAGCATTC 2651
 Db 1858 GGGATATGATGAGCCTATTAATTAAGATGGCATCTTTCAGAAAGTCTGGCAGCATTC 1917
 Qy 2652 TACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATATAAATGGGTTTTATTTCATAGA 2711
 Db 1918 TACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATATAAATGGGTTTTATTTCATAGA 1977
 Qy 2712 CATATTAACCTGATATATTTTTCATGCTGATGGTGCATATTAATTAATGACTGACTTT 2771
 Db 1978 CATATTAACCTGATATATTTTTCATGCTGATGGTGCATATTAATTAATGACTGACTTT 2037
 Qy 2772 GGCCTCTGCACTGGCTCAGATGAGACACAGATTTCAAGTACTATCAGAGTGGTGACCAT 2831
 Db 2038 GGCCTCTGCACTGGCTCAGATGAGACACAGATTTCAAGTACTATCAGAGTGGTGACCAT 2097
 Qy 2832 CCACGGCAAGATACATGATTTTCAGTAAATCAATGGGGGATCCCTCAAGCTGCTGATGT 2891
 Db 2098 CCACGGCAAGATACATGATTTTCAGTAAATCAATGGGGGATCCCTCAAGCTGCTGATGT 2157
 Qy 2892 GGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCAGCCAGCAGCAGCGATGCTAGCA 2951
 Db 2158 GGGACAGACTGAAGCCATTAGAGCGGAGAGCTGCAGCCAGCAGCAGCGATGCTAGCC 2217
 Qy 2952 CATCTTTGGTGGGACTCCCAATATATTCGACCTGAAGTGTGCTACGAAACAGGATAC 3011
 Db 2218 CATCTCTGTTGGGACTCCCAATATATTCGACCTGAAGTGTGCTACGAAACAGGATAT 2277
 Qy 3012 ACACAGTGTGATGCTGGTGGAGTGGTCTTATCTTTTGAATGTTGGTGGGACAA 3071
 Db 2278 ACACAGTGTGATGCTGGTGGAGTGGTCTTATCTTTTGAATGTTGGTGGGACAA 2337
 Qy 3072 CCTCCTTTCTTGGCACAACACCACTTAGAACACAAATGAAGGTTATCACTGGCAACA 3131
 Db 2338 CCTCCTTTCTTGGCACAACACCACTTAGAACACAAATGAAGGTTATCACTGGCAACT 2397
 Qy 3132 TCTCTTCACATTCACCAAGCTTAAGTCACTGATGCTGCTGATGAAATGAAGTCAAT 3191
 Db 2398 TCTCTTCACATTCACCAAGCTTAAGTCACTGATGCTGCTGATGAAATGAAGTCAAT 2457
 Qy 3192 CTTTGGCGAGACCGGAGATCGCTTAGGCAAGATGGTCTGATGAATGAAGTCAAT 3251
 Db 2458 CTGTGTCGAGCGAGGAGACCGCTCGGCAAGAACCGGTGCTGATGAGATGAAGGCTCAT 2517
 Qy 3252 CCATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGAGCTGCTGCTTCATACATT 3311
 Db 2518 CCATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGAGCTGCTGCTTCATACATC 2577
 Qy 3312 CCTAAATACACACCAACAGATACATCAATTTTGTATCTGTTGATCTGATGATTAATA 3371
 Db 2578 CCTAAATACAGCATCCCAACAGATACATCAATTTTGTATCTGTTGATCTGATGATTAATA 2637
 Qy 3372 TGGAGTATGATACGAGGAGGAAATGAATGACACTCTCAATGATGATGATTAATAAT 3431
 Db 2638 TGGAGTATGATACGAGGAGGAAATGAATGACACTCTCAATGATGATGATTAATAAT 2697

Qy 3432 GGAAGCATCTGAACATGCAATCTATGAATTTACCTTCGGAAGTCTTTTGTGATGACAA 3491
 Db 2698 GGAAGCATCTGAACATGCAATCTATGAATTTACCTTCGGAAGTCTTTTGTGATGACAA 2757
 Qy 3492 GGTACCCATATAATATTCGGAAGCTTATGAATATGAATATTAATTCACAGGCTCA 3551
 Db 2758 GGTACCCATATAATATTCGGAAGCTTATGAATATGAATATTAATTCACAGGCTCA 2817
 Qy 3552 GAGCAGAGCTGGATGAAGATGATCAAAACACAGCTCAGAGATTAATAATCGGATCTA 3611
 Db 2818 GAAACACAGCTGATGAAGATGATCAAAACACAGCTCAGAGATTAATAATCGGATCTA 2877
 Qy 3612 GTATATGTTTAAACACACTAGTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 3671
 Db 2878 GTGATGTTTAAACACACTAGTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 2933
 Qy 3672 CGAGGTGTTTGAAGTCTGAGAGTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 3731
 Db 2934 CAGGGTCTTGAAGTCTGAGAGTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 2985
 Qy 3732 GCTCTGTGACAAATATTTTATTTTCTTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 3791
 Db 2986 GCTCTGTGACAAATATTTTATTTTCTTAAATTAATGAATGAGGATTTGTAAGGCGCTGAAATG 3045
 Qy 3792 ATTCCAGCGCTTAAATCAGTATTTAGAAAAAATTTGTAAGGCGCTGAAATG 3851
 Db 3046 ATTCCAGCGCTTAAATCAGTATTTAGAAAAAATTTGTAAGGCGCTGAAATG 3105
 Qy 3852 CTGAATATATGATGATGCTGCTGATGCTTAAAGTCTTAAATTAATGAATGAGTAG 3900
 Db 3106 CTGAGTATATGATGATGCTGCTGATGCTTAAAGTCTTAAATTAATGAATGAGTAG 3154

RESULT 6

AA32981
 ID AAX32981 standard; DNA; 2442 BP.

XX AC AAX32981;

XX AC AAX32981;

DT 30-JUN-1999 (first entry)

XX Human warts protein encoding DNA.

XX Human; warts protein; Drosophila; protein kinase; epithelial cell;

XX cell regulation; cell proliferation; cell differentiation; tumour; ss.

XX Homo sapiens.

XX WO9915558-A1.

XX 01-APR-1999.

XX 24-AUG-1998; 98WO-JP03739.

XX 24-SEP-1997; 97JP-0258689.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;

XX WPI; 1999-244389/20.

XX P-PSDB; AAY06890.

XX Gene of vertebrates homologous to a Drosophila gene

XX Claim 7; Page 44-52; 68pp; Japanese.

XX The invention relates to a human warts protein, homologous to Drosophila
 CC warts gene expression product. The warts expression product is a protein
 CC kinase involved in the regulation of proliferation and differentiation of
 CC epithelial cells and suppression of tumour formation. The gene, its
 CC expression product and antibodies are useful in the study of the

CC mechanisms of tumour development. The present sequence represents a DNA
 XX encoding a human warts protein.
 SQ Sequence 2442 BP; 777 A; 528 C; 491 G; 646 T; 0 other;

Query Match 60.8%; Score 2422; DB 20; Length 2442;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2436; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 1062 TCTGGAACATGAATACGTAATCTCCCAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 1121
 DB 1 TCTGGAACATGAATACGTAATCTCCCAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 60
 QY 1122 GAGGCGTATCCCTCCACCACTCTCAACACTTCCCTCCATGAATCTCTCTCTCAATCAAGGACAG 1181
 DB 61 GAGGCGTATCCCTCCACCACTCTCAACACTTCCCTCCATGAATCTCTCTCTCAATCAAGGACAG 120
 QY 1182 AGAGGCAATAGTCTGTTCTGTTGGCAGACCAACCAATCATCATGCGAGAGTCTTAGCAAA 1241
 DB 121 AGAGGCAATAGTCTGTTCTGTTGGCAGACCAACCAATCATCATGCGAGAGTCTTAGCAAA 180
 QY 1242 TTTAACTTTCCATCAGGAGAGCTGGAATGCCAGATGGTACTGGCAAACTGATTTCTATG 1301
 DB 181 TTTAACTTTCCATCAGGAGAGCTGGAATGCCAGATGGTACTGGCAAACTGATTTCTATG 240
 QY 1302 ATACACCAAAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
 DB 241 ATACACCAAAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 1362 CTGACAGCAGCTAATGACAAAGCCCTTCTGCTTTTACAAACAGGGGGATCTGCTGCTCT 1421
 DB 301 CTGACAGCAGCTAATGACAAAGCCCTTCTGCTTTTACAAACAGGGGGATCTGCTGCTCT 360
 QY 1422 TCGTCATATACAAATGGAAGTATCTCTAGTCTATGATGCTGCTGCTGCTGCTGCTGCTAT 1481
 DB 361 TCGTCATATACAAATGGAAGTATCTCTAGTCTATGATGCTGCTGCTGCTGCTGCTGCTAT 420
 QY 1482 AACATGGAATATATAACATTAAGTGTACCTGGACTGCAAAACAAATGGGCTCAGTCACTCT 1541
 DB 421 AACATGGAATATATAACATTAAGTGTACCTGGACTGCAAAACAAATGGGCTCAGTCACTCT 480
 QY 1542 TCTGCTCCAGCCAGTCAATCCCGAGCAGTGGGATGGAATCCCTACATGGCAACCTAAC 1601
 DB 481 TCTGCTCCAGCCAGTCAATCCCGAGCAGTGGGATGGAATCCCTACATGGCAACCTAAC 540
 QY 1602 ATACAGTCAGCTCAAAATCTTTTAAATACCCATAGGAATAGCAAGTCACTCTGCT 1661
 DB 541 ATACAGTCAGCTCAAAATCTTTTAAATACCCATAGGAATAGCAAGTCACTCTGCT 600
 QY 1662 AATCTCAGCCCTTCTGCTACACAGTCACTGCAATACACAGCTCCTATTCAACAGCCT 1721
 DB 601 AATCTCAGCCCTTCTGCTACACAGTCACTGCAATACACAGCTCCTATTCAACAGCCT 660
 QY 1722 GTGAAAGATGCGGTATTAACACAGAGTACAGCTGCTTTAGCAGCTACACACCT 1781
 DB 661 GTGAAAGATGCGGTATTAACACAGAGTACAGCTGCTTTAGCAGCTACACACCT 720
 QY 1782 TCTTGATACACAGCAGCAATCAAACTGTTCAACCCAGTCTTTCTCTGAGGGAACCCGCT 1841
 DB 721 TCTTGATACACAGCAGCAATCAAACTGTTCAACCCAGTCTTTCTCTGAGGGAACCCGCT 780
 QY 1842 TCAATGTGACTGTGATGCGACCTGTTGCTGAAGCTTCCAACTATCAAGGACCCACCA 1901
 DB 781 TCAATGTGACTGTGATGCGACCTGTTGCTGAAGCTTCCAACTATCAAGGACCCACCA 840
 QY 1902 CCTTACCCAAAACATCTGCTGCAACCAACCCATCTGTTCTCCATACGAGTCAATCACT 1961
 DB 841 CCTTACCCAAAACATCTGCTGCAACCAACCCATCTGTTCTCCATACGAGTCAATCACT 900
 QY 1962 AAGCCTAGCAAGAGGATCAGCAGCTTGGCCAGGAAGATGAGAGTGAAGAGTTAT 2021
 DB 901 AAGCCTAGCAAGAGGATCAGCAGCTTGGCCAGGAAGATGAGAGTGAAGAGTTAT 960

QY 2022 GAAATGTTTATAGTGGGATAAAGAAAAGAACAGATTACAACTTCACTTACTGTT 2081
 DB 961 GAAATGTTTATAGTGGGATAAAGAAAAGAACAGATTACAACTTCACTTACTGTT 1020
 QY 2082 AGGAAAAACAAGAGATGAAGAGCGAAGGAAATCTCGTATTCAAAGTTATTCTCTCTCAA 2141
 DB 1021 AGGAAAAACAAGAGATGAAGAGCGAAGGAAATCTCGTATTCAAAGTTATTCTCTCTCAA 1080
 QY 2142 GCATTTAAATCTTTATGAGGACACATGTAGAAAATGTACTCAATCTCATCAGAGCGT 2201
 DB 1081 GCATTTAAATCTTTATGAGGACACATGTAGAAAATGTACTCAATCTCATCAGAGCGT 1140
 QY 2202 CTACATCGTAAAAAACAATTAGAGATGAATGATCGGGTGGATTATCTCAAGATGCC 2261
 DB 1141 CTACATCGTAAAAAACAATTAGAGATGAATGATCGGGTGGATTATCTCAAGATGCC 1200
 QY 2262 CAGGATCAATGAGAAAGATGCTTTGCCAAAAGAAATCTAAATFACATCGTCTTAAAGG 2321
 DB 1201 CAGGATCAATGAGAAAGATGCTTTGCCAAAAGAAATCTAAATFACATCGTCTTAAAGG 1260
 QY 2322 GCTAAATCGCAAGCTCTATGTTGTGAAGATAAAGACACTAGGAATAGAGCAATTTGCT 2381
 DB 1261 GCTAAATCGCAAGCTCTATGTTGTGAAGATAAAGACACTAGGAATAGAGCAATTTGCT 1320
 QY 2382 GAAGTCTGTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCCAAACAAACTCTTCCA 2441
 DB 1321 GAAGTCTGTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCCAAACAAACTCTTCCA 1380
 QY 2442 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTTAAAGCTGAGAGATATCTCTG 2501
 DB 1381 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTTAAAGCTGAGAGATATCTCTG 1440
 QY 2502 GCTGAAGCTGACAACTAAGTGGTGTCTCTATATTATTCATTCACAAATAGGCAAT 2561
 DB 1441 GCTGAAGCTGACAACTAAGTGGTGTCTCTATATTATTCATTCACAAATAGGCAAT 1500
 QY 2562 TTATATCTTTGAATGAGCTACATTTCTGGGGGTGATATGATAGCTATTAATTTAGAATG 2621
 DB 1501 TTATATCTTTGAATGAGCTACATTTCTGGGGGTGATATGATAGCTATTAATTTAGAATG 1560
 QY 2622 GGCATCTTTCCAGAAAGTCTGGCAGATTTCTACATAGCAGAACTTACCTGTGAGTTGAA 2681
 DB 1561 GGCATCTTTCCAGAAAGTCTGGCAGATTTCTACATAGCAGAACTTACCTGTGAGTTGAA 1620
 QY 2682 AGTGTTCATAAAATGGTTTTATTCATAGATATTAACACCTCATATAATTTGATTGAT 2741
 DB 1621 AGTGTTCATAAAATGGTTTTATTCATAGATATTAACACCTCATATAATTTGATTGAT 1680
 QY 2742 CGTGATGGTTCATATAATTAATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2801
 DB 1681 CGTGATGGTTCATATAATTAATTTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
 QY 2802 GATTTCTAAGTACTATCAGAGTGGTACCATCCACGGCAAGATAGCATGATTTCACTAAT 2861
 DB 1741 GATTTCTAAGTACTATCAGAGTGGTACCATCCACGGCAAGATAGCATGATTTCACTAAT 1800
 QY 2862 GAATGGGGATCCCTCAAGCTGTGATGAGAGCAGACTGAAGCCATTAGAGCGGGA 2921
 DB 1801 GAATGGGGATCCCTCAAGCTGTGATGAGAGCAGACTGAAGCCATTAGAGCGGGA 1860
 QY 2922 GCTGACGCGCAGCAGCAGGATGCTTAGCAGATTTCTTGGTGGGACTCCCAATATAT 2981
 DB 1861 GCTGACGCGCAGCAGCAGGATGCTTAGCAGATTTCTTGGTGGGACTCCCAATATAT 1920
 QY 2982 GCACCTGAAAGTGTGCTAGCAAGATACACAGTGTGATGTTGGTGGAGTGTGTTG 3041
 DB 1921 GCACCTGAAAGTGTGCTAGCAAGATACACAGTGTGATGTTGGTGGAGTGTGTTG 1980
 QY 3042 GTTATCTTTTGAATGTTGGTGGGACAACTCTTCTTGGCACAACCACTTAGAA 3101
 DB 1981 GTTATCTTTTGAATGTTGGTGGGACAACTCTTCTTGGCACAACCACTTAGAA 2040
 QY 3102 ACACAAATGAAGTATCAACTGGCAACATCTCTTACATTCACCAACCAAGCTAACTC 3161

XX	Human warts gene fragment encoding a partial peptide fragment.
DE	
XX	
XX	Human; warts protein; Drosophila; protein kinase; epithelial cell;
KW	cell regulation; cell proliferation; cell differentiation; tumour; ss.
KW	
XX	
OS	Homo sapiens.
XX	
XX	WO9915558-A1.
PN	
XX	
XX	01-APR-1999.
PD	
XX	
XX	24-AUG-1998; 98WO-JP03739.
PF	
XX	
XX	24-SEP-1997; 97JP-0258689.
PR	
XX	
XX	(SUNE) SUMITOMO ELECTRIC IND CO.
PA	
PI	Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;
PI	
DR	WPI; 1999-244389/20.
DR	P-PSDB; AAY06891.
XX	
XX	Gene of vertebrates homologous to a Drosophila gene
PT	
XX	
PS	Disclosure; Page 52-56; 68pp; Japanese.
XX	
CC	The invention relates to a human warts protein, homologous to Drosophila
CC	warts gene expression product. The warts expression product is a protein
CC	kinase involved in the regulation of proliferation and differentiation
CC	epithelial cells and suppression of tumour formation. The gene, its
CC	expression product and antibodies are useful in the study of the
CC	mechanisms of tumour development. The present sequence represents a
CC	human warts gene fragment encoding a partial peptide fragment.
XX	
XX	Sequence 1374 BP; 442 A; 252 C; 288 G; 392 T; 0 other;
SQ	
	34.1%; Score 1360.4; DB 20; Length 1374;
	Query Match

```
Query Match
34.18; Score 1360.4; DB 20; Length 1374;
```

Seq	Sequence	1357 BP	436 A	254 C	285 G	382 T	0 other
	Query Match	31.0%	Score 1233.2	DB 22	Length 1357		
	Best Local Similarity	97.3%	Pred. No. 1.1e-298				
	Matches 1321	Conservative	0	Mismatches 23	Indels 14	Gaps 6	
QY	2276	AAAGATGCTTTGGCAAAAAGATCTAATTACATCCGCTTTAAAGAGGCTAAATGGACAA	2335				
Db	1	AAAGATGCTTTGGC-AAAGAAATCTAATTACATCCGCTTTAAAGAGGCTAAATGGACAA	59				
QY	2336	GTCTATCTTTTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTCTAGC	2395				
Db	60	GTCTATCTTTTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTCTAGC	119				
QY	2396	AAGAAAGTAGACTAAGGCTTTGTATGCAACAAAAAAGCTCTCGAAAGAAAGATGTTCT	2455				
Db	120	AAGAAAGTAGACTAAGGCTTTGTATGCAACAAAAAAGCTCTCGAAAGAAAGATGTTCT	179				
QY	2456	TCCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAAGCTGACAA	2515				
Db	180	TCCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAAGCTGACAA	239				
QY	2516	TGAATGGGTAGTTCGCTATATATTATTCATTCCAAGATAAGACAAATTTATCTTTGTAAT	2575				
Db	240	TGAATGGGTAGTTCGCTATATATTATTCATTCCAAGATAAGACAAATTTATCTTTGTAAT	299				
QY	2576	GGACTACATTCCTGGGGTGCATATGATAGCCATTATTAATAGATGGCATCTTCCAGA	2635				
Db	300	GGACTACATTCCTGGGGTGCATATGATAGCCATTATTAATAGATGGCATCTTCCAGA	359				
QY	2636	AAGTCTGGCAGCATTCACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAAT	2695				
Db	360	AAGTCTGGCAGCATTCACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAAT	419				
QY	2696	GGGTTTTATTCATAGAGATATTAACCTGATAATATTTTGATGTATCGTGATGTCATAT	2755				
Db	420	GGGTTTTATTCATAGAGATATTAACCTGATAATATTTTGATGTATCGTGATGTCATAT	479				

QY 2756 TAAATGACTGACTTTGGCTTCTGCACTGGCTTCAGATGGACACACGATTTCTAAGTACTA 2815
 DB 480 TAAATGACTGACTTTGGCTTCTGCACTGGCTTCAGATGGACACACGATTTCTAAGTACTA 539
 QY 2816 TCAGATGGTGCACCATCCAGGCAAGATAGATGATTTTCAGTAATGAATGGGGGATCC 2875
 DB 540 TCAGATGGTGCACCATCCAGGCAAGATAGATGATTTTCAGTAATGAATGGGGGATCC 599
 QY 2876 CTCAGCTGTCGATGTGGAGACAGCTGAAGCCATTAGAGCGAGAGCTGCACGCCAGCA 2935
 DB 600 CTCAGCTGTCGATGTGGAGACAGCTGAAGCCATTAGAGCGAGAGCTGCACGCCAGCA 659
 QY 2936 CCAGCGATGCTAGCACATTTTGGTGGGACTCCCAATTATATTCACCTGAAGTGT 2995
 DB 660 CCAGCGATGCTAGCACATTTTGGTGGGACTCCCAATTATATTCACCTGAAGTGT 719
 QY 2996 GCTACGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTTGA 3055
 DB 720 GCTACGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTTGA 779
 QY 3056 AATGTTGGTGGGACAACCTCTTTTGGCACAAACACCATTAGAAACAAATGAAGGT 3115
 DB 780 AATGTTGGTGGGACAACCTCTTTTGGCACAAACACCATTAGAAACAAATGAAGGT 839
 QY 3116 TATCAACTGCCAACATCTTTCACATTCACACAAAGCTAACTCAGTCTGAAGCTTC 3175
 DB 840 TATCAACTGCCAACATCTTTCACATTCACACAAAGCTAACTCAGTCTGAAGCTTC 899
 QY 3176 TGATCTATTATTAATTAATTTTGGAGGACCGAAGATGCTTAGCAAGAATGGTCTGA 3235
 DB 900 TGATCTATTATTAATTAATTTTGGAGGACCGAAGATGCTTAGCAAGAATGGTCTGA 959
 QY 3236 TGAATTAAGCTATCCATTTTAAAAA---CAATTGACTTCT-CCAGTGAACCTGA--- 3288
 DB 960 TGAATTAAGCTATCCATTTTAAAAA---CAATTGACTTCTCCAGTGAACCTGA--- 1019
 QY 3289 ---GACAGAGCTGCTTATCATATTCCTTAAATCACAC-ACCAACAGATACATCAA 3343
 DB 1020 CAGCAGCTGGCTTTCAACAATTCCTTAAATCACAC-ACCAACAGATACATCAA 1079
 QY 3344 TTTTGATCC-TGTTGATCCGATAAATATGAGTGATGATAACGAGGAAGAAATGTA 3402
 DB 1080 TTTTGATCCGATAAATATGAGTGATGATAACGAGGAAGAAATGTA 1139
 QY 3403 ATGACACTCTCAATGGATGATAAATGGAAGCATCTGACATGATATGAT 3462
 DB 1140 ATGACACTCTCAATGGATGATAAATGGAAGCATCTGACATGATATGAT 1199
 QY 3463 TTACCTTCCGAAGGTTTTTGTGATGACATGGCTACCCATATAATTATCCGAAGCTATTG 3522
 DB 1200 TTACCTTCCGAAGGTTTTTGTGATGACATGGCTACCCATATAATTATCCGAAGCTATTG 1259
 QY 3523 ATATGATATACATTAATTCACAGGCTCAGAGCAGCTGCGATGAAGATGATCAAAACA 3582
 DB 1260 ATATGATATACATTAATTCACAGGCTCAGAGCAGCTGCGATGAAGATGATCAAAACA 1319
 QY 3583 CAGGCTCAGAGATTAATAATCGCATCTAGTATATGTT 3620
 DB 1320 CAGGCTCAGAGATTAATAATCGCATCTAGTATATGTT 1357

RESULT 9

AAX32984 standard; DNA; 795 BP.

XX AAX32984;

AC AAX32984;

DT 30-JUN-1999 (first entry)

DE Human warts gene fragment encoding a partial peptide fragment.

XX Human; warts protein; Drosophila; protein kinase; epithelial cell;

KW cell regulation; cell proliferation; cell differentiation; tumour; ss.
 XX Homo sapiens.
 OS WO9915558-A1.
 PN 01-APR-1999.
 PD 24-AUG-1998; 98WO-JP03739.
 PF 24-SEP-1997; 97JP-0258689.
 PR (SUME) SUMITOMO ELECTRIC IND CO.
 XX Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;
 PI WPI; 1999-244389/20.
 DR P-PSDB; AAY06893.
 DT Gene of vertebrates homologous to a Drosophila gene
 PT Disclosure; Page 58-60; 68pp; Japanese.
 PS The invention relates to a human warts protein, homologous to Drosophila
 XX warts gene expression product. The warts expression product is a protein
 CC kinase involved in the regulation of proliferation and differentiation of
 CC epithelial cells and suppression of tumour formation. The gene, its
 CC expression product and antibodies are useful in the study of the
 CC mechanisms of tumour development. The present sequence represents a
 CC human warts gene fragment encoding a partial peptide fragment.
 XX Sequence 795 BP; 224 A; 228 C; 160 G; 183 T; 0 other;
 SQ

Query Match 19.88; Score 788.6; DB 20; Length 795;
 Best Local Similarity 99.5%; Pred. No. 2.4e-187;
 Matches 791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 786 GTTCTCTCAGAGGCGATGCCCGCCACTAGGAGAAAGTGTGGCTATCATCTCTGAGAGTCC 845
 DB 1 GTTCTCTCAGAGGCGATGCCCGCCACTAGGAGAAAGTGTGGCTATCATCTCTGAGAGTCC 60
 QY 846 AACTCACAGACAGATGAGGAAGACCTTTCTGTGATCTGTATATCAGCATTTGTTCAA 905
 DB 61 AACTCACAGACAGATGAGGAAGACCTTTCTGTGATCTGTATATCAGCATTTGTTCAA 120
 QY 906 GCTCACCTTAGCAAGGACAGAGAGTGAACCCGCCACCCACCTCAAGTAAGAGGTGT 965
 DB 121 GCTCACCTTAGCAAGGACAGAGAGTGAACCCGCCACCCACCTCAAGTAAGAGGTGT 180
 QY 966 ACTCTCTCCACCTCCCAAGAGGCCAGACTCCCGCTCCCAAGAGGTACAACTCCACCTCC 1025
 DB 181 ACTCTCTCCACCTCCCAAGAGGCCAGACTCCCGCTCCCAAGAGGTACAACTCCACCTCC 240
 QY 1026 CCTTCATGGGAACCAACTCTCAACAAAGCGTATCTGGAACATGGAATACGTAATC 1085
 DB 241 CCTTCATGGGAACCAACTCTCAACAAAGCGTATCTGGAACATGGAATACGTAATC 300
 QY 1086 TCCGGAATCTCTCTCTCCACCTGGGGCATGGCAAGAGGCTATCTCCACACCTCTC 1145
 DB 301 TCCGGAATCTCTCTCTCCACCTGGGGCATGGCAAGAGGCTATCTCCACACCTCTC 360
 QY 1146 AACACTTCCCGCATGAATCTCTCAATCAAGGACAGAGGCGATTAGTTCTCTCTCTGT 1205
 DB 361 AACACTTCCCGCATGAATCTCTCAATCAAGGACAGAGGCGATTAGTTCTCTCTCTGT 420
 QY 1206 GGCAGACAACCAATCATCATGAGAGTCTTACCAATTTAACTTTCCATCAGGGAGACT 1265
 DB 421 GGCAGACAACCAATCATCATGAGAGTCTTACCAATTTAACTTTCCATCAGGGAGACT 480
 QY 1266 GGAATGCAAGTGGTACTGGACAAACTGATTTTCATGATACACCAAAATGTTCTCCTGCT 1325
 DB 481 GGAATGCAAGTGGTACTGGACAAACTGATTTTCATGATACACCAAAATGTTCTCCTGCT 540

QY 1326 GGCACGTGTAATGCGGAGGACACACCTCCATATCTCTGACAGCAGCTAATGGACAAAGC 1385
 Db 541 GGCACGTGTAATGCGGAGGACACACCTCCATATCTCTGACAGCAGCTAATGGACAAAGC 600
 QY 1386 CTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTCTGCTATATACAAATGGAAGTATT 1445
 Db 601 CTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTCTGCTATATACAAATGGAAGTATT 660
 QY 1446 CCTCAGCTATGATGTGCGCAACAGAAATAGTACATAAGTGAACATATATACATTAGT 1505
 Db 661 CCTCAGCTATGATGTGCGCAACAGAAATAGTACATAAGTGAACATATATACATTAGT 720
 QY 1506 GTACCTGGAGCTGCAACAAATGGGCTCAGTCAATCTTCTGCTCCAGCCAGTATCCCCG 1565
 Db 721 GTACCTGGAGCTGCAACAAATGGGCTCAGTCAATCTTCTGCTCCAGCCAGTATCCCCG 780
 QY 1566 AGCAGTGGGCATCAA 1580
 Db 781 AGCAGTGGGCATCAA 795
 RESULT 10
 AAT42120
 ID AAT42120 standard; cDNA; 3155 BP.
 AC AAT42120;
 XX
 DT 31-JAN-1997 (first entry)
 XX
 DE M-lats2 gene encoding large tumour suppressor.
 XX
 KW Mouse; m-lats2 gene; large tumour suppressor; fetal brain;
 KW protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnery; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX
 OS Mus musculus.
 XX
 FH Key
 FT CDS
 FT 1..2943
 FT /*tag= a
 FT /product= m-lats2 protein
 XX
 PN W09630402-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 26-MAR-1996; 96WO-US04101.
 XX
 PR 27-MAR-1995; 95US-0411111.
 XX
 PA (UYVA) UNIV YALE.
 XX
 PI Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX
 DR WPI: 1996-455275/45.
 XX
 DR P-PSDB; AAW06287.
 XX
 PT New isolated large tumour suppressor gene - used to develop prods.
 XX for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Claim 8; Page|133-137; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats2
 CC protein, and has been isolated from a newborn mouse brain phage
 CC lambda-ZAP cDNA library using a 2.2-kb DNA probe from the Drosophila
 CC lats gene (AAT42117). A homologous mouse sequence has also been
 CC isolated (m-lats, AAT42119). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other

CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;
 Query Match 19.1%; Score 760; DB 17; Length 3155;
 Best Local Similarity 63.7%; Pred No. 6.8e-180;
 Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;
 QY 1653 CACTCTGCTTAATCTCAGCCTTCTGCTACACAGTCACTGCAATTTACACCAGCTCTCTATT 1712
 Db 985 CAACCTGAGCCCTCACTGCGCGCCCAACACGGTCACCGCTGACGCCGCCACACATC 1044
 QY 1713 CAACAGCCTGTGAAGAATGATGGTGTATTAACACAGAGCTACAGACTGCTTTAGCACCT 1772
 Db 1045 CTTACCCCTGTGAAGAGCGTGGTGTGCGCGCCGAGCCGACAGACGCCGTGGGGCC 1104
 QY 1773 ACACACCCCTTCTGGATACACAGCAATTCAAACTGTTTCAACCCAGTCTCTTTCTCTGAG 1832
 Db 1105 TCGCACCCCTGCTGGTGGCTGCGCCACAGCACCTGCCACTGAGAGCTGGAGAGGAG 1164
 QY 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGA 1892
 Db 1165 GAGGCGAGCGCGCCACACCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGC 1224
 QY 1893 CCACACACACCTTACCCAAACATCTGTGACACCAACCCATCTGTTCTCCATACGAG 1952
 Db 1225 CCACCGCTTCCGTATCCAAAGCACTTGTGCTGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
 QY 1953 TCAATCAGTAAGCTAGCAAGAGGATCAACCAAGTCTGCCCAAGCAAGTACAGAGTGA 2012
 Db 1285 GACCTGGACAGCTGTGACACAGTGTGACAGAGTCTCGAGGGGGGCACTGCTTAGAC 1344
 QY 2013 AAGAGTTA-----TGAAATGTTGATAGTGGGGATAAAGAAAGAAACAG 2057
 Db 1345 GGGAGTGAACAAGACCAACAGGTGCGAAGGAGACAAAGCTGGCAGACACAAAGCAG 1404
 QY 2058 ATTACAACTTCACTTACTGTTAGGAAAACAGAAAGATGAGAGCGAAGGAAATCT 2117
 Db 1405 ATTACAGCTCCCGTGGCTGTGCTGCCCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 QY 2118 CGTATTCAAAGTATTCTCTCAAGCATTTAAATCTTTATGAGCAACATCTAGAAAT 2177
 Db 1465 CGATCAAGAGTACTCCCTTTATGCTTCAATCTTCTGAGCAACACCTGGGAGAAAT 1524
 QY 2178 GTACTCAATCTCATCAGCAGGCTCTACATCGTAAAAACAATTAGAGAAATGAAATGATG 2237
 Db 1525 GTCATCAAAACCTACCAGCAGAGGTGACCGCGAGGCTACAGCTGCGAGCAAAATGGCC 1584
 QY 2238 CGGTTTGGATTATCTCAAGATCCCGAGGATCAATAGAGAAAGATGCTTTGCCAAAGAA 2297
 Db 1585 AAAGCTGGCTCTGTGAGGCCGAGCAGAGTGAAGAGATGCTCTTACCAGAGGAG 1644
 QY 2298 TCTAATTACATCGTCTTAAAGGGCTAAATGCGAAGTATGTTTGTGAAGATAAG 2357
 Db 1645 TCTAATAACACCGCTGAAGAGGCCAAGATGGCAAGTCCATGTTGTGAAATCAAG 1704
 QY 2358 ACACCTAGGAATAGGAGCAATTTGGTGAAGTCTGCTAGCAAGAAAGATAGATACTAAGCT 2417
 Db 1705 ACTTAGGATCGGTGCTTGGGGAAGTGTGCTCGCTTGAAGCTGAGCACTCACGCT 1764
 QY 2418 TTGTATGCAAAACAACTCTTCAAGAAAGATGTTCTTCTGAAATCAAGTCTCAT 2477
 Db 1765 CTGTAGCCCATGAAGACTCTCAGGAAGAGGATGCTTGAACCCGGAATCAAGTGGCCAT 1824
 QY 2478 GTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACAATGAATGGTACTGCTCTATAT 2537
 Db 1825 GTCAGGCTGAGAGGAGACATCTGCTGAAGCAGACAAATGAGTGGTGTGCTCAACTAC 1884
 QY 2538 TATTCAATTCGAAGATAAGGACAATTTATATCTTTGTAATGAGTACTACATTTCTGGGGTGTAT 2597

Db 1885 TACTCCCTTCCAGGACAGGACGCTGCTACTTTGTGTGACTACATACACAGGGGGAT 1944
 Qy 2598 ATGATGAGCCTATTAAATAGAAATGGGATCTTTCCACAAAAGTCTGGACGATTTCTACATA 2657
 Db 1945 ATGATGAGCCTGCTGATCAGATGGAGGTCTCCCTGAGCAGCCTGGCCGCTTCTACATT 2004
 Qy 2658 GCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGTGTATTATTCATACAGATATT 2717
 Db 2005 GCAGATTGACCTGGCCGCAATTAAGAGTGTCCACAAAGTGGCTTTATTCACCGGGACATC 2064
 Qy 2718 AAACCTGATATATTGTTGATGTCGAGTGTGATATTAATTAAGTACTGACTTTGGCCCTC 2777
 Db 2065 AAGCCTGACAACTACTCATCGACTGGATGGTCATATTAAAGCTGACAGATTGTCCTC 2124
 Qy 2778 TGCAGTGGCTTTCAGATGACACAGCAGTCTTAAGTACTATCAGAGTGTGACCATCCACGG 2837
 Db 2125 TGCAGTGGATTCAGTGGAGCTCAATTCCTCAAGTACTACAGAAAGGGAACACATGAGA 2184
 Qy 2838 CAAGATAGCATGATTTCAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCGCTGACCTCTGGGACGATGTTTCCAACGTGCTGTGGAGAC 2244
 Qy 2898 AGACTGAAGCATTAGACGGAGAGTGTGACGCCACACAGGATGTCTAGACATTTCT 2957
 Db 2245 AGTTAAAGACCTGGAGGACAGGGGCGCAGAAAGCAGACAGAGGTGCTGGCACATTTCT 2304
 Qy 2958 TTGTTGGGACTCCCAATATATATGACCTGAACTGAGTGTCTGACGAACAGATACACAG 3017
 Db 2305 CTGTCGGGACACCAATATACATCGCTCCGGAGGTCTCTCGCAAGAGGTACACGAG 2364
 Qy 3018 TTCTGTGATGTGGAGTGTGGTGTATCTTTTGAATGTTGGTGGGACAACTCTCT 3077
 Db 2365 CTCTGTGACTGTGGAGCTGGTGTGATTTCTTTGAGATGCTGTGGGACGCGCT 2424
 Qy 3078 TTCTTGGGCAACAAACCATTTAGAACACAAATGAAGTTATCACTGGCAACATCTCTT 3137
 Db 2425 TTCTTGGGCCCCACCCACAGAGAGCGAGCTGAAGGTGATCACTGGGAGACGCTG 2484
 Qy 3138 CACATTCACCAACAGCTAACTAGTCTCTGAGCTTCTGATCTTATTATTAACTTTGC 3197
 Db 2485 CATATCTCTTACGAGGTGAGGCTCAGCGCTGAGGCGCGAGACCTCATCACGAGCTGTGC 2544
 Qy 3198 CGAGGACCGGAGATCGCTTAGCAAGATGCTGCTGATGAATAAAAGCTCATCCATT 3257
 Db 2545 TCGCGGCTGACTGCGCGCTGGCAGGATGGGCGAGATGACCTCAAGGCGACACCCGTT 2604
 Qy 3258 TTTAAACAATTTGACTTCTCCAGTACCTGAGACAGCAGTCTGCTTATCATATTCTCTAA 3317
 Db 2605 TTCAACACCATGCTGCTTTTCCGCTGACATCGAAGACAGGCTGCACCTACCTCCACC 2664
 Qy 3318 ATCACACACCCACAGATACATCAATTTTGTATCTGTGATCTGTGATTAATTTAGGT 3377
 Db 2665 ATCAGCCACCCCATGGACACCTTCCAAATTTTACCGGTGATGAAGAAAGCCCTGGCAC 2724
 Qy 3378 GATGATACGAGGAGAAATTAATGACACTCTCAATGATGGTATATAAATGGAAG 3437
 Db 2725 GAGGCCAGC---GGAGAGCGCCAGGCCCTGGACAGCTGGCCCTCCCGCAGCAGCA 2781
 Qy 3438 CATCTGACATGCTATGATTAATTTACCTTCCGAAGGTTTTTTTGTATGACAAATGGCTAC 3497
 Db 2782 CATCAGAGCAGCGCTTCTATGAGTTACCTTCCGAGGTTCTTCGATGACAGCGGCTAT 2841
 Qy 3498 CCATATATTATCCGAGCC 3517
 Db 2842 CCCTCCGCTGCCGAAGCC 2861

RESULT 11
 AA251507
 ID AA251507 standard; DNA; 3155 BP.
 XX
 AC AA251507;

XX
 DT 21-JUN-2000 (first entry)
 XX
 DE Mouse Lats2 (large tumour suppressor) DNA.
 XX
 KW Mouse; Lats2; large tumour suppressor; cytostatic; vulnery;
 KW cell-overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
 KW treatment; prevention; screening; cancer; skin; ovarian tumour;
 KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
 KW LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
 KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
 KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.
 XX
 OS Mus musculus.
 XX
 FH Location/Qualifiers
 FT 1..2943
 FT /*tag= a
 FT /product= "Lats2 protein"
 FT 970..1920
 FT /*tag= b
 FT /note= "This region is erroneously repeated in the
 mouse Lats2 DNA sequence shown in figure.14"
 FT 1921..2960
 FT /*tag= c
 FT /note= "This region is missing in the mouse Lats2
 DNA sequence shown in figure 14"
 XX
 WO200010602-A1.
 XX
 PD 02-MAR-2000.
 XX
 PF 18-AUG-1999; 99WO-US19068.
 XX
 PR 18-AUG-1998; 98US-0096996.
 PR 18-AUG-1998; 98US-0096997.
 XX
 PA (UYA) UNIV YALE.
 XX
 PI Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
 PI Turenchalk GS, Stewart RA;
 XX
 DR WPI: 2000-246496/21.
 DR P-PSDB; AAY70392.
 XX
 PT Use of lats proteins, complexes of lats and cdc2 for treating cancer
 PT that is refractory to treatment by standard chemotherapy and radiation
 PT therapy, and disorders associated with aberrant levels of cdc2 activity
 XX
 PS Claim 44; Pages 112-117; 134pp; English.
 XX
 CC The present sequence is a DNA encoding mouse Lats2 (large tumour
 CC suppressor) protein which is a cell overproliferation inhibitor and a
 CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 CC The present sequence is useful for treating cancer that is refractory
 CC to standard chemotherapy or radiation therapy such as hyperplasia,
 CC metaplasia, or dysplasia, and disorders associated with aberrant
 CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
 CC include degenerative disorders, growth deficiencies, hypoproliferative
 CC disorders, physical trauma, lesions, and wounds. An animal model
 CC preferably a mouse, in which a lats gene has been disrupted by homologous
 CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX
 SQ Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;
 Query Match 19.1%; Score 760; DB 21; Length 3155;
 Best Local Similarity 63.7%; Pred. No. 6.8e-180;
 Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 1653 CACTCTGCTAAATTCAGCCTTCTGCTACACAGTCACTGCAATTAACACAGTCTCTAT 1712
 Db 985 CAACCTGAGCCCTACTGCGGCCCAACACAGGTCAACCGGTGAGCGCCGACACATC 1044
 QY 1713 CAACAGCTGTGAAGATGCGGTATTAACACAGAGTCAAGACTGCTTTAGCACCT 1772
 Db 1045 CTTACCCCTGTGAAGAGCGTGTGCTGCGGCCGAGCCGACAGACCGCTGGGGCC 1104
 QY 1773 ACACACCCCTCTGGATACACAGCAATTCACCTGTTCAACCCAGTCTTTTCTGTAG 1832
 Db 1105 TCGCACCCCGCTGGGTGCTGCGGCCACAGCACCTGCGCACTGAGAGCGCTGGAGAG 1164
 QY 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTGCTGTAAGCTCCAACTATCAAGA 1892
 Db 1165 GAGGCGAGCGAGCCACACCCGCTGGATGTGGACTATGGCGGCTCCGAGCGCAGTGC 1224
 QY 1893 CCACACACCTCAACCAACATCTGTGACCAACCAACCTCTCTCTCCATACAGAG 1952
 Db 1225 CCACCGCTTCCGTATCCAAAGCACTTGTCTGCCAGTAAAGTCTGAGCAGTACAGCGTG 1284
 QY 1953 TCATCAGTAAGCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATCAGAGTGA 2012
 Db 1285 GACCTGGACAGCTGTGCACAGTGTGCAGCAGAGTCTGCGAGGGGGCACTGATAGAC 1344
 QY 2013 AAGAGTTA-----TGAATAATGTTGATAGTGGGATAAAGAAAGAAACAG 2057
 Db 1345 GGGAGTGACAGAGCCACAAAGTGGAGGAGGACAAAGCTGGCAGACAAAAGCAG 1404
 QY 2058 ATTCAACCTTCACTTACTGTTAGGAAACAAAGAGATGAAGAGCGAAGGAATCT 2117
 Db 1405 ATTACAGCTTCCCGTGTCTGCCGAAGATACAGAGATGAAGAGAGAGAGTCT 1464
 QY 2118 CGTATTCAGATTTCTCCTCAAGCATTTAAATCTTTATGGACCAACATGTAGAAAT 2177
 Db 1465 CGCATCAGAGTTACTCCCTTTATCCCTTCAAAATCTTCATGGAGCAACACGTGGAGAT 1524
 QY 2178 GTACTCAATCTCATCAGCAGCGCTTACATCGTAAAAAACAATTAGAATGAATGATG 2237
 Db 1525 GTCATCAAAACCTACCAGCAAGAGTCCAGCGGAGGTACAGCTGGAGCGAATGGCC 1584
 QY 2238 CGGGTGGATATCAAGATCCGAGGATCAAAATGAGAAAGATGCTTTGCGCAAAAGAA 2297
 Db 1585 AAAGCTGGCTCTGTGAGCGCCGAGGAGCAGATGAGGAAGATCCTTACCAGAAGAG 1644
 QY 2298 TCTAATACATCCGCTTAAAGGGCTAAATGGACAGTCTATCTTTGTGAAGATAAG 2357
 Db 1645 TCTAATCAACCGCTGAAGAGGCGCAAGATGGACAGTCCATGTTGTGAATCAAG 1704
 QY 2358 ACACATAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATCTAAGGCT 2417
 Db 1705 ACTCTAGGCATCGTGCCTTTGGGGAAGTGTGCTCGCTTGTGAAGTGGACACTCACGCT 1764
 QY 2418 TTGTATGCAAAAACCTCTTCCGAAGAAGATGTTCTTCTCGAAATCAAGTCCGCTCAT 2477
 Db 1765 CTGTAGCGCATGAAGACTCTCAGGAAGAGGATGCTCTGAACCGGAATCAAGTGGCCCAT 1824
 QY 2478 GTTAAGGCTGAGAGATATCTGCTGAGCTGACATGAATGGGTAGTTCGTCTATAT 2537
 Db 1825 GTCAAGCTGAGAGGAGATCTGCTGAAGCAGACAAATGAGTGGGTGCTCAAACTCTAC 1884
 QY 2538 TATTCATTCGAAGATAAGCAATTTATCTTGTGAAGTACATCTCTGGGGGTGAT 2597
 Db 1885 TACTCTTCCAGACAGGACAGCCTGTACTTTGTGTGAGTACATACACCGCGGGAT 1944
 QY 2598 ATGATGAGCCTATTAATAGATGGGCATCTTTCCAGAAAGTCTGGCAGATTTACATA 2657
 Db 1945 ATGATGAGCCTGCTGATCAGGATGGAGTCTTCCCTGAGCACCTGGCCGCTTCTACATT 2004
 QY 2658 GCAGAACTTACCTGTGCAGTTCAAGATGTTTCAATAAATGGGTTTATTCATAGAGATAT 2717
 Db 2005 GCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAAGATGGGCTTTATCCACCGGACATC 2064

QY 2718 AAACCTGATATATTTTGAATGATGCTGATGCTCATATTAATTAATGACTGACTTTGGCCTC 2777
 Db 2065 AAGCCTGACAACTACTCATCGACCTGGATGCTCATATTAAGCTGACAGATTTTGGCCTC 2124
 QY 2778 TGCATGGCTTCAGATGGACACAGGATTAAGTACTATCATGAGTGGTGGACATCCACGG 2837
 Db 2125 TGCATGGATTCAGGTGGACTCACAATTCGAAGTACTACACAGAAAGGAAACCATGAGA 2184
 QY 2838 CAAGATAGCATGATTCAGTAATGAATGGGGGGATCCCTCAAGCTGCTGATGTGGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAAGTGTGCTGGAGAC 2244
 QY 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCAGCAGCAGCTGCTAGCATATCT 2957
 Db 2245 AGTTTAAGACCTGGAGCAGAGGGCGGAGAGCAGCAGCAGGTGCTGCGCATCTCT 2304
 QY 2958 TTGGTGGGACTCCCAATTAATTCACCTGAAAGTGTGCTACGAACAGGATACACAG 3017
 Db 2305 CTGTGCGGACACCAAAATTCATCGTCCGAGGTGCTTCCGCAAAAGGTTACACGAG 2364
 QY 3018 TTGTGCTGATGCTGAGTGTGCTGTTATCTTTTGAATGTTGGTGGGACAACTCTCT 3077
 Db 2365 CTCTGTGACTGTGTGAGGCTGCTGTGATCTCTTTGAGATGCTGTTGGGCGCGCT 2424
 QY 3078 TTTCTTGGCACAACACCACTTAGAAAACAAATGAAGGTATCAACTGGCAACATCTCTT 3137
 Db 2425 TTTCTTGGCCCCCACCACAGAGAGCGCTGAGAGTGATCACTGGGAGAGCAGCTG 2484
 QY 3138 CACATTTCCACCAAGCTAAACTCAGTCTGAAAGTCTTCTGATCTTATTAATTAACCTTTGC 3197
 Db 2485 CATATCCCTTACGAGTGTGAGGCTCAGCGCTGAGGCCGAGACCTCATCAGGAAGCTGTGC 2544
 QY 3198 CGAGGACCGAAGATCGCTTAGGCAAGATGCTGATGAATGAATGAAGCTCATCATTT 3257
 Db 2545 TCGCGGCTGACTGCGGCTGGGCGAGGATGGGCGAGATGACCTCAAGSCACACCCGTTTC 2604
 QY 3258 TTTAAACAAATGACTTCTCCAGTGCACCTGACACAGCAGTCTGCTTATACATTTCTTAAA 3317
 Db 2605 TTCAACACCATCGACTTTCCGCTGCATCCGAAGCAGGCTGACCTTACCTGCCACC 2664
 QY 3318 ATCACAACCCAAAGATACATCAAAATTTGATCTGTTGATCTGCTGATAAATATTTGAGT 3377
 Db 2665 ATCAGCCACCCCATGGACACCTTCCAAATTTGACCCGGTGGATGAAGAAAGCCCTGGCAC 2724
 QY 3378 GATGATACGAGGAAGAAATGAATGACACTCTCAATGGATGCTATAAATGAAG 3437
 Db 2725 GAGGCCAGC---GGAGAGAGCCCAAGGCTGGGACAGCTGGCCCTCCCCCAGCAGCAG 2781
 QY 3438 CATCTGACATGCTATCTATGAATTTACCTTCCGAAGGTTTTTTTGTGACAAATGGCTAC 3497
 Db 2782 CATCCAGAGCAGCGCTTCTATGAGTTCACCTTCCGAGGTTCTTCGATGACAAAGCTAT 2841
 QY 3498 CCATATAATTTCCGAAGCC 3517
 Db 2842 CCTTCCGGTGCCCGAAGCC 2861

RESULT 12

AA87397
 ID AA87397 standard; cDNA; 5276 BP.

XX AA87397;

XX 08-OCT-1999 (first entry)

XX Human WART2 cDNA.

XX WART2; hWART2; WART orthologue; human; signal transduction;
 KW protein kinase; cancer; tumour; diagnosis; therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

5'UTR 1..374
 CDS /*tag= a
 375..3641
 3'UTR /*tag= b
 3642..5276
 /*tag= c

WO9937787-A2.
 29-JUL-1999.
 20-JAN-1999; 99WO-US01145.
 21-JAN-1998; 98US-0072023.
 (SUGEN-) SUGEN INC.
 Flanagan P, Plowman GD;
 WPI; 1999-458698/38.
 P-PSDB; AAY06527.
 New nucleic acid encoding human orthologs of Drosophila WART proteins, used to identify specific modulators for treating cancer or for diagnosis
 Claim 1; Page 122-123; 137pp; English.
 This is the nucleotide sequence of a cDNA clone coding for a human orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor serine/threonine kinase WART. hWART2 cDNA was isolated from a human bone marrow cDNA library using a PCR fragment of hWART1 as probe. hWART2 is consistently expressed in human tumour cells lines, except for most of the colon cancer lines examined. This overexpression in tumour cells versus normal tissues may provide a target for oncology drug development. Nucleic acids encoding full-length hWART2 and hWART2 polypeptides lacking one or more of amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518, and 974-1048, or lacking one or more of the N-terminal domain, catalytic domain, or C-terminal domain are claimed, as well as hWART1 sequences (AAX87396), hWART2 and hWART1 polypeptides, antibodies, a method for identifying modulators of hWART function, and use of such modulator compounds to treat an abnormal condition involving hWART signal transduction, especially cancer. Probes for detection of hWART nucleic acids are also claimed.
 Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;
 Query Match 18.7%; Score 744.4; DB 20; Length 5276;
 Best Local Similarity 66.3%; Pred. No. 7e-176;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;
 QY 1872 GAAGCTCCAACTATCAAGGACACACACACCTACCAAAACATCTGCT-----GCACC 1926
 DB 1893 GGAGGCCCCAGACGGAGGTGCCGCCCTCCGCCCTACCGAAGCACCTGCTGCTGCGAGC 1952
 QY 1927 AAAACCCATCTGTTCTCCATACAGAGTCAATCACTAGAGCTAGCAAGAGGATCAG-----1982
 DB 1953 AAGTCGAGCAGTAGACCTGGAGAGCCTGTGCGCAGGATGGAGAGAGCCTCCGTGCG 2012
 QY 1983 ---CCAAAGCTTGGCCAAAGGAGATGAGAGTCAAAAGAGTTATGAAATGTTGATAGTGGG 2039
 DB 2013 GGCCCCAACGAGCCGAGGGGGGCGACAAAGAGCGGCAAGAGGGGCAAGAGGC 2072
 QY 2040 GATTAAGAAAGAAACAGATTACACTTACCTATTACTGTAGGAAACACAGAAAGAT 2099
 DB 2073 GGAAGGATATAAAGACAGATTACAGACCTCTCCCGTTCCTCCGCAAAACACAGCAGAGAC 2132
 QY 2100 GAAGAGCAAGGGAATCTCTGATTCAAGTATTCTCTCAAGCATTTAAATCTTTATG 2159
 DB 2133 GAAGAGAGAGAGTAGTCAAGCATCAAGAGCTACTCGGCATACGCCCTTAAGTCTTCATG 2192
 QY 2160 GAGCAACATGTAGAAATGTACTCAAAATCTCATCAGCAGGCTCTATCATGTAACAAACAA 2219

DB 2193 GAGCAGCAGCTGGAGAAATGTCTATCAAAACCTTACCAGCAGAAAGGTTAAACCGGAGCTCGAG 2252
 QY 2220 TTACAGATGAATGATGATGCGGTTGGATTATCTCAAGATGCCAGGATCAATAGAGAAAG 2279
 DB 2253 CTGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGAGCAGATCGGAAG 2312
 QY 2280 ATGCTTTGCCAAAGAAATCTAATTTACATCGTCTTAAAGGGCTAAATGGCAAGTCT 2339
 DB 2313 ATCTCTTACCAGAAAGAGTCTAATTTACACAGGTTAAAGAGGGCCAGATGGACAGTCT 2372
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGA 2399
 DB 2373 ATGTTTGTCAAGATCAAAACCTCGGATCGCTTTGGAGAAGTGTGCTTGTCTTGT 2432
 QY 2400 AAGTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAGAAAGATGTTCTTCT 2459
 DB 2433 AAGTGGACACTCAGGCCCTGTACGCCATGAAGACCTTAAGAAAAAGGATGTCTGAAC 2492
 QY 2460 CGAAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTTGGCTGAAGCTGACAATGAA 2519
 DB 2493 CGGAATCAGTGGGCCACGTCAGGCCGAGAGGACATCTTGGCCGAGGACAGCAATGAG 2552
 QY 2520 TGGTAGTTCGTCTATATTATTCATTCAGAGATAAGGACAAATTTATCTTTGTAATGGAC 2579
 DB 2553 TGGGTGTCAAACTCTACTCTCTCTCAAGACAAAGACAGCCCTGTACTTTGTGATGGAC 2612
 QY 2580 TACATTCCTGGGGTGATATGATGAGGCTTATTAATTTAGATGGGCATCTTTCCAGAAAGT 2639
 DB 2613 TACATCTCTGGTGGGACATGATGAGCTGTGATCGGATGGAGGTCTTCCCTGAGCAC 2672
 QY 2640 CTGCGACGATTCTACATAGCAGAACTTACCTGTGTCAGTGTGAAAGTGTTCATAAATGGT 2699
 DB 2673 CTGCCCCGTTCTACATCGCAGAGCTGACTTGGCCATTTGAGAGTGTCCAAAGATGGC 2732
 QY 2700 TTTATTCATAGATATTAACCTGATATTTTGGATTTGATGTCGTGATGTCATATTAAA 2759
 DB 2733 TTTATTCACCCAGACATCAAGCCCTGATAACATTTTGTATGATGTCGTGATGTCATATAA 2792
 QY 2760 TTGACTACATTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2819
 DB 2793 CTCACAGATTTGGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2852
 QY 2820 AGTGGTCAACCATCACCGCAAGATAGCATGGATTTCAATGAATGGGGGATCCCTCA 2879
 DB 2853 AAGGGAGCCATGTACAGAGGACAGCATGGAGCCAGCCAGCCCTGCGGATGATGTCT 2912
 QY 2880 AGCTGTCGATGTGGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACCCAGCAGCAG 2939
 DB 2913 AACTGTGCGTGTGGGACAGGCTGAAGACCTTAGAGCAGAGGGCGGGAAGCAGCAGCAG 2972
 QY 2940 CGATGCTACACATCTTTGGTGGGACTCCCAATTTATTTGACCTGAAGTGTGCTA 2999
 DB 2973 AGTGGCTGACATCTTCTGTTGGGACTCCAACTTACATCGCAGCCGAGGTGCTCTC 3032
 QY 3000 CGAACAGATTACACAGTGTGTGATTGGTGGAGTCTGGTGTATTCTTTTGAAGATG 3059
 DB 3033 CGAAAGGTTACACTCAACTCTGTGACTGTTGGAGTGTGGAGTGTCTTCTCGAGATG 3092
 QY 3060 TTGGTGGGCAACCTCTCTTTTGGCAACAAACACTTATAGAAACACAAATGAAGTTATC 3119
 DB 3093 CTGGTGGGCGAGCGCCCTTTTGGCACCTTACTCCACACAAACCCAGCTGAAGGTGATC 3152
 QY 3120 AACTGGCAACATCTCTTACATTTCCACACAGTCAACTAGTCTCTGAGCTTCTGAT 3179
 DB 3153 AACTGGGAGAACCGCTTCCACATTTCCAGCCAGGTGAAGCTGAGCCCTGAGCCGAGGAC 3212
 QY 3180 CTATTATTAACTTTGCCCGAGGACCGGAAGATCGCTTAGGCAAGATGTCGTGATGAA 3239
 DB 3213 CTATCACCAAGCTGTGCTCTCGCAGACACCGCTGGGGCGGATGGGGCGGATGAC 3272
 QY 3240 ATAAAGCTCATCTTTTAAACAAATTTGACTTCTTCCAGTGAACCTGAGACAGAGTCT 3299

Db 3273 CTGAAGCCGCCCCCTTTCTTCAAGCCCAATGACTTCTCCAGTGACATCCGGAAGCAGCCA 3332
 QY 3300 GCTTATACATATCTCTAAATACACACCCCAACAGATACATCAAAATTTTGATCTGTGTAT 3359
 Db 3333 GCCCCTACGTTCCACCATACGCCACCCCATGACACCTCGGAATTCGACCCCGTAGAT 3392
 QY 3360 CTTGATAAATTTATGGAGTGATGATAAGGAGGAGAGAAATGTAATGACACTCTCTATGGA 3419
 Db 3393 GAAGAAAGCCCTTGGACGATGCCAGC---GAAGGTAGCACCAGGCGCTGGGACACACTC 3449
 QY 3420 TGTATATAAATGAAGATCTGACATGCAATCTATGCAATTTACCTTCCGAAGTTT 3479
 Db 3450 ACCTGCCCAATACAGCATCTGACGACGCAATTTTACGAATTCACCTTCCGAAGTTT 3509
 QY 3480 TTTGATGACATGCTACCCATATATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3539
 Db 3510 TTTGATGACATGCTACCCCTTCGATGCCAAGCCTTCAGGAGCAGGAAGCTTCACAG 3569
 QY 3540 TCACAAAGGCTCAGA 3553
 Db 3570 GCTGAGAGCTCAGA 3583

RESULT 13

AA59129
 ID AA59129 standard; DNA; 5486 BP.

AC AA59129;

DT 07-NOV-2000 (first entry)

XX DNA encoding a tumour suppressor protein hGHITS1.

XX Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 387..3653
 FT /tag= a
 FT /product= "tumour suppressor protein"

XX EP1022333-Al:

XX 26-JUL-2000.

XX 07-OCT-1999; 99BP-0119199.

XX 25-JAN-1999; 99JP-0016223.

XX (JCRP-) JCR PHARM CO LTD.

XX Koga J, Kono K, Zolotar'ov FN;

XX WPI: 2000-516013/47.

XX P-PSDB; AAB07633.

XX New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 PT diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 PT nephropathy or cardiopathy.

XX Claim 1; Page 18-28; 59pp; English.

XX The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnostic pharmaceutical preparations can be used for examining
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant

CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.
 XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;
 Query Match 18.7%; Score 744.4; DB 21; Length 5486;
 Best Local Similarity 66.3%; Pred No. 7.2e-176;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;
 QY 1872 GAAGCTCCAACTATCAAGGACCAACACCTTACCACCAAAACATCTGCT-----GCAC 1926
 Db 1905 GGAGGCCAGACCGGAGGTGCCCGCTCCCGCTACCGAAGCACCTGCTGCTGCGCAGC 1964
 QY 1927 AAACCCATCTGTTCTCCATACGAGTCAATCAGTAGCCTAGCAAGAGGATCAG---- 1982
 Db 1965 AGTCGGAGCAGTACGACGCTTGGCAGGCTGTGCGAGGCTGGAGCAGAGCTCCCTGGG 2024
 QY 1983 ---CCAAAGTTTGCCCAAGGAAGATGAGAGTGAAGAGTTATGAAATTTGTGATGCTGG 2039
 Db 2025 GGCCCAACGAGCCGCGGCGGCGGACAGAGCCGCAAGGCCCAAGGGGACAAAGC 2084
 QY 2040 GATAAAGAAAAGAACAGATTTACAACTTCACTTACTTCTTAGGAAAACAAAGAAAGAT 2099
 Db 2085 GGAAAGGATAAAAGCAGATTCAGACCTCTCCCGTTCCCGTCCGCAAAACACAGCAGAC 2144
 QY 2100 GAAGCGGAGGAATCTCGTATTCAAGTTATTTCTTCAAGCATTTAAATTTCTTTATG 2159
 Db 2145 GAAGAGAGAGAGAGTCAAGCATCAAGAGCTACTCGCCATACGCTTAAAGTTCTTCATG 2204
 QY 2160 GAGCAACATGTAGAAAATGTACTCAAACTCTATCAGCAGCGTCTACATCGTAAAAACAA 2219
 Db 2205 GAGCACAGCTGGAGAAATGTATCAAAACCTACCAGAGAGTTAACCGGAGGTGCAG 2264
 QY 2220 TTAGAAATCAAAATGATGCGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAG 2279
 Db 2265 CTGGACCAAGAAATGCCAAAGCTGACCTCTGTGAAGCTGAGCAGGAGCAGATGCGGAAG 2324
 QY 2280 ATGCTTTGCCAAAAGAAATCTAATTTACATCCGCTTTTAAAGGCTTAAATGACAACTCT 2339
 Db 2325 ATCTCTACCAAGAAAGAGTCTAATTTACAAGAGTTAAAGAGGCGCAAGATGGAAGTCT 2384
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGATAGGAGCAATTTGGTGAAGCTGTCTAGCAAGA 2399
 Db 2385 ATGTTTGTCAAGATCAAAACCTTGGGATCGGTGCGCTTTGGAGAGTGTGCTTGTGT 2444
 QY 2400 AAAGTAGATTAAGGCTTTGTATGCAACAAAACCTTTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 2445 AAGGTGGACACTCACGCCCTGTACGCCATGAAGACCCCTAAGGAAAAAGATGTCTCTGAAC 2504
 QY 2460 CGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACATGAA 2519
 Db 2505 CGGAATCAGGTGGCCCGCACGTCAGGCGGAGGAGATCCTGGCCGAGGACACAAATAG 2564
 QY 2520 TGGTAGTTCGCTCTATATTATTATTCATTCACAGATAAGGACAAATTTATCTTTGTAATGGAC 2579
 Db 2565 TGGTGTGTCAAACTCTACTACTCTTCCAAAGACAAAGACAGCCTGTACTTGTGATGGAC 2624
 QY 2580 TACATTCCTGGGGGTGATATGATGAGCCTTATTAATTTAGATGGGCACTTTTCCAGAAAT 2639
 Db 2625 TACATCCCTGGTGGGACATGATGAGCCTGTGATCCCGATGGAGTCTTCCCTGAGCAC 2684
 QY 2640 CTGGCAGCATTTACATAGCAGACACTTACCTGTGCAAGTTGAAGTGTTCATAAATGGGT 2699
 Db 2685 CTGGCCCGGTTTACATCCGAGAGTGAAGTTTGGCCATTTGAGAGTGTCCACAGATGGGC 2744
 QY 2700 TTATTATAGAGATATTAACCTGATATATTTTATTGATGCTGATGCTGATATTAATAA 2759
 Db 2745 TTCATCCCGGAGACATCAAGCCTGATACATTTTGTATAGATCTGATGCTGATTAATAA 2804
 QY 2760 TTGACTGACTTTGGCCCTCTGCACTGGCTTCAGATGAGACACAGATTTCTTAAGTACTATCAG 2819

Db 2805 CTCACAGTTTCGGCTCTGCTCACTGGTTTCAGGTGACATCAAAATTCACCAATATACACAG 2864
 QY 2820 AGTGGTACACATCCAGCGGCAAGATAGCATGATTTTCAGTAATGAATGGGGGATCCCTCA 2879
 Db 2865 AAAGGGAGCCATGTCAGACAGGACAGCATGGAGCCCGACCTCTGGGATGATGTCT 2924
 QY 2880 AGCTGTGATGGACAGACATGAAGCCATTAGACGGGAGAGCTCCAGCCGACACACAG 2939
 Db 2925 AACTGTGCTGGTGGGACAGGCTGAAGACCTAGACAGAGGGCGGGAAGCAGCAGCAG 2984
 QY 2940 CGATGTCTAGCAGATTTCTTGTGGAGTCTCCCAATATATATTCACCTCAAGTGTGCTA 2999
 Db 2985 AGTGTGCTGACATTTCACTGTGGGAGTCTCCAAATATATATTCACCTCAAGTGTGCTA 3044
 QY 3000 CGAACAGGATACACAGATTTGTGATTTGTGGAGTGTGGTGTATTTTGAATG 3059
 Db 3045 CGCAAGGTACACTCACTCTGTGAGTGTGGAGTGTGGAGTGTCTCTTCGAGATG 3104
 QY 3060 TTGGTGGGACACCTCTTTTGGCACAAACACCAATATAGAACAACAAATGAAGTTATC 3119
 Db 3105 CTGTTGGGCGAGCGGCTTTTGGCACCTACTTCCACACAGAAACCCAGCTGAAGTGTATC 3164
 QY 3120 AACTGCAAAACATCTCTTCCATTCACACAGCTTAACTCAGTCTCAAGCTTCTGAT 3179
 Db 3165 AACTGGGAGACAGCTCCACATTCAGCCCGAGGTGAAGCTGAGCCCTGAGCCGAGGAC 3224
 QY 3180 CTTATTATTAATTTTCGCGAGGACCCGAAAGATCGCTTAGGCAAGAATGCTGCTGATGA 3239
 Db 3225 CTCATCAACCAAGCTGTGCTCCGACAGACACCGCTGGGCGGAATGGGCGCATGAC 3284
 QY 3240 ATAAAGCTCATCCATTTTAAACAAATGACTTCTCCAGTGTGCTGAGACAGCAGTCT 3299
 Db 3285 CTGAAGGCGCCACCTCTTCTAGCGCCATTTGACTTCTCCAGTGTGACTCCGGAAGCAGCA 3344
 QY 3300 GCTTATCATATCTTAAATACACACCAACAGATACATCAATTTGATCTCTGAT 3359
 Db 3345 GCCCTAGTTCACCATCAGCCACCCCTGAGACCTCGAATTTTCAGCCCGTAGAT 3404
 QY 3360 CTGTATAATTAAGGATGATGATACAGGAGGAAGAAATGTAAATGACACTCTCAATGGA 3419
 Db 3405 GAAGAAGCCCTTGAACGATGCCAGC---GAAGTAGCACCAAGGCTGGGACACATC 3461
 QY 3420 TGGTATAAATGGAAGCATCTCAACATGCTATGATTTTACCTTCGGAAGTTT 3479
 Db 3462 ACCTCGCCCAATGAAGCATCTGAGCAGCATTTTACGAATTCACCTTCGGAAGTTT 3521
 QY 3480 TTTGATGACATGGCTAGCCATATATATTCGAAGCCCTATTGAATTAATGATACATTAAT 3539
 Db 3522 TTTGATGACATGGCTAGCCCTTTTCGATGCCCAAGCCCTTCAGGAGCAGAGCTTCACAG 3581
 QY 3540 TCACAAGGCTCAGA 3553
 Db 3582 GCTGAGAGCTCAGA 3595

RESULT 14

AA59130
 ID AA59130 standard; DNA; 5486 BP.

XX
 AC
 XX
 DT
 XX

07-NOV-2000 (first entry)

DE DNA encoding a tumour suppressor protein hGHITS2.

XX Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiodysplasia;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.

OS Homo sapiens.

XX key Location/Qualifiers

FT

387..3653
 /*tag= a
 /product= "tumour suppressor protein"

EP1022333-A1.

26-JUL-2000.

07-OCT-1999; 99EP-0119199.

25-JAN-1999; 99JP-0016223.

(JCRP-) JCR PHARM CO. LTD.

Koga J; Kono K; Zolotaryov FN;

WPI; 2000-516013/47.

P-PSDB; AAB07664.

New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 diagnosis of dwarfism, gigantism, acromegaly, angiodysplasia, diabetic
 nephropathy or cardiopathy

Claim 1; Page 28-39; 59pp; English.

The present sequence encodes a human growth hormone inhibited tumour
 suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 antineoplastic pharmaceutical preparation. Probes for the hGHITS
 DNA sequences can be used in diagnostic pharmaceutical preparations.
 The diagnostic pharmaceutical preparations can be used for examining
 expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 angiodysplasia, diabetic nephropathy or cardiopathy; or in malignant
 tumours including breast cancer, renal adenocarcinoma, colorectal
 cancer, and leukaemia. Antibodies against the proteins can be used in
 a diagnostic pharmaceutical preparation for examining expression of
 a tumour suppressor gene.

Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;

Query Match 18.7%; Score 744.4; DB 21; Length 5486;
 Best Local Similarity 66.3%; Pred. No. 7.2e-176;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;

QY 1872. GAAGCTCAAACTATCAAGGACCAACCCCTACCAAAACATCTGCT-----GCACC 1926
 Db 1905. GGAGGCCACGACCGAGGTGCCCCCTCCGCGCTTACCGAGACCTGCTCTCGCAGC 1964
 QY 1927. AAAACCATCTGTTCTCCATACAGATCAATCAGTAAGCTTAGCAAGAGATGAG----- 1982
 Db 1965. AAGTCGGAGCAGTACGACCTGCGACAGCTGTGCGAGCATGGAGAGCCTCCGTGCG 2024
 QY 1983. ---CCAAAGCTTGGCCCAAGGAGATGAGATGAAAGAGTTATGAAATGTTAGTGGG 2039
 Db 2025. GGCCCAACAGCCGAGGCGGCGGACAGAGCCGCAAGAGCCGCAAGGGGCAAGAGG 2084
 QY 2040. GATAAAGAAAAGAACAGATTACAACTTACCTATTACTGTAGGAAAACAAAGAGAT 2099
 Db 2085. GGAAGGATAAAAAGAGAGATTACAGACCTCTCCCGTTCCCGCAAAACAGCAGAGAC 2144
 QY 2100. GAAGAGCAAGGGGATCTCGTATTCAAGTATTCTCTCAAGCATTTAAATCTTTATG 2159
 Db 2145. GAAGAGAGAGAGAGTACCGCATCAAGAGCTACTCGCATACCGCTTTAAGTTCTCATG 2204
 QY 2160. GAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCTCTACATCTTAAACAA 2219
 Db 2205. GAGCAGACCTGGAGATGTATCAAAACCTACAGCAGAGGTTAACCGAGGCTCGAG 2264
 QY 2220. TTAGAGAAATGAATGATGCGGGTTGGATTATCTCAAGATCCCGAGGATCAATAGAGAG 2279
 Db 2265. CTGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTCAGCAGGAGCAGATGCGGAG 2324
 QY 2280. ATGCTTTGCCAAAAGAGATCTAATACATCGTCTTAAAGGGCTTAAATGGCAAGTCT 2339

Db 2325 ATCCCTACAGAAAGAGTCTAATTAACAAGGTTAAAGAGGCCAAGATGGACAAGTCT 2384
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGCTTAGCAAGA 2399
 Db 2385 ATGTTTGTGAAGATCAAAACCCCTGGGGATCGGTGCGCTTTGGAGAAGTGGCCCTTGTGT 2444
 QY 2400 AAGTATGATCTAAGGCTTTGATGCAACAAAACCTTTCCGAAGAAGAGATGTTCTTCTT 2459
 Db 2445 AAGTGGACACTCACGCCCTGTACGCCATGAAGACCCCTAAGGAAAAGAGATGCTCTGAAC 2504
 QY 2460 CGAAATCAAGTCCCTCATGTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACATGAA 2519
 Db 2505 CGGAATCAGGTGGCCCAAGCGGAGGAGGAGATCTGCGCCGAGGACAGCAATAG 2564
 QY 2520 TGGGTAGTTCGTCTATATTATTCATTCACAAAGATAAGACAAATTTATCTTTGTAATGGAC 2579
 Db 2565 TGGGTGTCNAACCTACTACTCTCTCCRAAGACAAGACAGCTGTACTTTGTGATGGAC 2624
 QY 2580 TACATTCCTGGGGGTGATATGATGAGCCTATTAATAGAAATGGGCATCTTTCCAGAAAGT 2639
 Db 2625 TACATCCCTGGTGGGACATGATGAGCCTCTCTGATCCGATGGAGCTCTTCCCTGAGCAC 2684
 QY 2640 CTGGCAGGATTCATACATAGACAGAACTTACCTGTGACAGTTGAAGTCTTATATAAATGGT 2699
 Db 2685 CTGGCCCGTCTTACATCGCAGAGCTGACTTTGGCCATTTAGAGAGTCTCCACAAGATGGC 2744
 QY 2700 TTATTTATAGAGATATTAACCTGATATATTTTATGATGCTGATGCTCATATATAA 2759
 Db 2745 TTTATCCAGGACATCAAGCTGATACATTTTGTATAGTCTGATGCTGATGCTCATATAA 2804
 QY 2760 TTGACTGATTTGGCCCTCTGCACTGCTCTGATGACACACAGATTTCTAAGTACTATCAG 2819
 Db 2805 CTACAGATTTTCGCCCTCTGCACTGGTTCAGCTGAGCTCACAATTTCCAAATATTACCAG 2864
 QY 2820 AGTGTGACCATCCAGGCAAGATACATGCTTTTCAAGTAAATGAGTGGGGATCCCTCA 2879
 Db 2865 AAGGGAGCCATCTAGACAGGACGATGAGCCAGCCAGCTCTGGGATGATGTGTCT 2924
 QY 2880 AGTGTGATGTGGACAGACTGAAGCCATTTAGAGCGGAGCTGCACGCCAGCAGCACAG 2939
 Db 2925 AACTGTGGTGGGACAGCTGAGACCCCTAGACAGGCGCGGAGCAGCAGCACAG 2984
 QY 2940 CGATGCTAGACATCTTTTGGTGGGACTCCCAATTAATTGCACTGAAAGTGTGCTA 2999
 Db 2985 AGTGTGCTGACATTTCACTGTGGGACTCCAAACTACATGCAACCCGAGGTGCTCTC 3044
 QY 3000 CGAACAGGATACACAGATGTTGTGATGTTGGTGGAGTGTGTTTCTTTTGAAGT 3059
 Db 3045 CGAAAGGTACACTCAACTCTGTGACTGTGGAGTGTGGAGTGTATCTCTCGAGATG 3104
 QY 3060 TTGGTGGGACAACCTCTTTTGGCACAACACCACTTAGAACACAAATGAAGGTATC 3119
 Db 3105 CTGGTGGGACGCCCTTTTGGCACCCTACTCCACAGAACCCAGCTGAAGTGTATC 3164
 QY 3120 AACTGCAACATCTTCCATATTCACACCAAGCTTAAGCTTAACTGCTCTGAAAGCTTCTGAT 3179
 Db 3165 AACTGGGCAACACAGCTTCCATATTCAGCCAGGTGAAGCTGAGCCTCAGGCCAGGGAC 3224
 QY 3180 CTTATTTAACTTTGGCGAGGCCGAGATCGCTTAGGCAAGAAATGGTGTGATGAA 3239
 Db 3225 CTCATACCAAGCTGTGCTGCTCCGAGACACCCGCGGAGTGGGCGGAGTGGCTGATGAC 3284
 QY 3240 ATAAAGCTATCCATTTTTTAAACAAATTTCTCCAGTGACCTGAGACAGAGCTCT 3299
 Db 3285 CTGAAGCCCAACCTCTTTCAGGCCATTTGACTTCTCAGTGACATCCGGAAGCAGCCA 3344
 QY 3300 GTTTCATACATCTCTAAATACACACCAACCAAGATACATCAAAATTTTGTGCTGTGAT 3359
 Db 3345 GCCCCTCTGTTCCCAACCATCAGCCACCCCTGAGACACCTCGAATTTCCGACCCCGTAGAT 3404
 QY 3360 CTGTATAAATTTGAGTGTATGATACGAGGAGAAATGTAATGACACTCTCAATGGA 3419
 Db 3405 GAAGAAGCCCTTGGAAACGATGCCAGC---GAAGGTAGCACCAAGGCTTGGACACACTC 3461

QY 3420 TGTATATAAATGGAAGCATCTCTGAACATGCTATGAAATTTACCTTTCCGAAGGTTT 3479
 Db 3462 ACCTCCCCCAATAACAGCATCTCTGAGCAGCATTTTAGAATTCACCTTCCGAAGGTTT 3521
 QY 3480 TTTGATGACAATGGCTTACCCATATATAATTTCCGAAGCCTTATTGAATATGAAATACATTAAT 3539
 Db 3522 TTTGATGACAATGGCTTACCCCTTTTCATGATGCCAAAGCCTTTCAGGACGAGAAGTTCACAG 3581
 QY 3540 TCACAAGGCTCAGA 3553
 Db 3582 GCTGAGAGCTCAGA 3595

RESULT 15
 AA261160
 ID AA261160 standard; cDNA; 1961 BP.
 XX AC AA261160;
 XX AC 30-MAY-2000 (first entry)
 XX CDNA SS1771A encoding a partial protein kinase.
 DE Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 3..1769 /tag= a
 FT /note= "partial sequence"
 XX WO200008180-A2.
 XX 17-FEB-2000.
 XX 03-AUG-1999; 99WO-US17630.
 XX 04-AUG-1998; 98US-0095270.
 XX 11-SEP-1998; 98US-0099972.
 XX (IMV) IMMUNEX CORP.
 PI Virca GD, Bird TA, Anderson DM, Marken JS;
 XX WPI; 2000-195584/17.
 DR P-PSDB; AAY69162.
 XX New human kinase polypeptides and polynucleotides used as molecular weight markers and as controls for peptide fragmentation

Claim 1; Page 9; 60pp; English.

The present sequence encodes a partial polypeptide which has kinase activity. The kinase polynucleotides can be used to express the polypeptides, and as probes to identify nucleic acids encoding proteins having kinase activity. The kinase polypeptides and fragmented polypeptides are used as molecular weight and isoelectric focusing markers, and as controls for peptide fragmentation. They also have a number of therapeutic uses as kinases play a central role in cellular signal transduction. The polypeptides could also be used to identify binding partner proteins. The polypeptides can also be used as a reagent to identify any proteins that the polypeptide regulates, and proteins with which it might interact. The polypeptides may also be used for preparation of antibodies. The antibodies can be used in assays to detect the presence of the protein, and to purify the protein by immunoaffinity chromatography.

Sequence 1961 BP; 504 A; 550 C; 539 G; 368 T; 0 other;

Query Match 18.6%; Score 742.8; DB 21; Length 1961;

Best Local Similarity 66.2%; Pred. No. 1.le-175;
Matches 1122; Conservative 0; Mismatches 557; Indels 15; Gaps 3;

1872	QY	GAAGTCTC	AAACTAT	CATC	AGGAC	CAC	CAAC	CCCTT	ACCC	AAAC	AT	TGCT	----	GCAC	1920					
21	Db	GGAGCC	CAGAC	CGGAG	GTG	CCGCC	CT	CGCC	CT	ACCG	AAAG	CACT	TGCT	GTGCGCAGC	80					
1927	QY	AAAA	CCCAT	CTGT	CTCT	CCAT	CAC	GAT	CAAT	CACT	TAAG	CCCT	TAC	AAAGAGGAT	CAG	--- 1982				
81	Db	AGT	CGGAC	AGT	AGC	CT	GGC	AG	CCCT	GTG	CC	AG	GCAT	TGAGCAG	AGCC	TCCCTGGG 140				
1983	QY	----	CCA	AGT	TG	CCCA	AGGA	AGAT	GA	GAG	AGT	GAAA	AGG	ATT	TAT	GAAAAT	TTGTGATGGG 2039			
141	Db	GGCC	CAAC	AGCG	CGCG	GGG	GGG	GGC	AC	AAAG	AGCG	CGCA	AAAA	CGCC	AAAG	GGGGACA	AAAGC 200			
2040	QY	GATA	AAAA	AAAA	GAAC	AGAT	TAC	AACT	TTC	ACCT	TA	TACT	CTT	AGG	AAAA	AAACA	AGAAAT 2099			
201	Db	GGA	AGGAT	AAAA	AGC	AGAT	TTC	AGAC	CCCT	CTC	CGCT	TTC	CGG	TCC	GAAA	AAAC	AGCAGAG 260			
2100	QY	GAAG	AGCA	AGG	GAAT	CTG	TA	TTC	AAAG	TTAT	TCT	CCT	CAAG	CAT	TAAAT	TCTTTATG 2159				
261	Db	GAAG	AGA	AGAG	AGAT	CAC	GCAT	CA	AGAG	CTACT	TG	CCAT	AC	GCCT	T	TAA	TTTCTTCATG 320			
2160	QY	GAG	CAAC	ATG	TAG	AAAT	TG	TACT	CA	AAAT	CT	ATC	AG	CGCT	T	ACAT	CGTAAAA	AAACA 2219		
321	Db	GAG	CAG	CAGT	TGG	AGAT	TG	TCT	CA	AAAC	CTC	AG	CGA	AGG	TTT	AC	CGGAGG	CTGCAG 380		
2220	QY	TTAG	AGA	AT	GAAT	GAT	CG	GGT	TGG	ATT	AT	CT	CA	AGAT	CCCC	AGAT	CAAA	TGAGAAAG 2279		
381	Db	CTG	GAG	CA	GA	AAAT	TGG	CA	AAAG	CTT	CT	GT	G	AGCT	GAG	CAG	CAGAT	TGCGGAAG 440		
2280	QY	ATG	CTTT	TCC	AAAA	AGAT	CT	AA	TTT	ACAT	CCG	CTT	TAA	AGG	CGCT	TAA	AT	TGAC	AGTCT 2339	
441	Db	ATC	CT	T	ACC	AGA	AG	AGT	CT	AA	TTT	AC	AG	GGT	TAA	AG	GGCC	AAAGAT	TGAC	AGTCT 500
2340	QY	ATG	TTT	TG	TAAG	ATA	AA	GAC	ACT	AG	GAAT	TAG	AG	CAT	TTT	GGT	GA	AGT	CTCT	

1041	Db	AACTGTCGGTGGGGACAGGCTGAAGACCCTAGAGCAGAGGGCGCGGAAGCAGCACCAG	1101
2940	QY	CGATGCTAGCACATCTTTTGGTTGGGACTCCCAATATATATGACACCTGAAGTGTGCTA	2999
1101	Db	AGGTGCTGGCACATCTACTGTGGGACTCCAACACTACATCGCACCCGAGGTGCTCCTC	1160
3000	QY	CGAACAGGATACACACAGTGTGTGATTTGGTGGAGTGTGGTGTATCTTTTGTGAAGT	3059
1161	Db	CGCAAAGGATACACATCAACTCTGTACTGTGGAGTGTGGAGTGAATCTCTTCGAGATG	1220
3060	QY	TTGGTTGGGACAACTCCCTTTCTTGGCACAAACACCATATAGAAACACAAAATGAAGGTTATC	3119
1221	Db	CTGGTGGGCGCCGCCCTTTTGGCACCTACTCCACAGAAACCCAGCTTGAAGGTGATC	1280
3120	QY	AACTGGCAAAACATCTCTTCACTATCCACCAACAGCTAAACTCACTCCTGAAGCTTCTGAT	3179
1281	Db	AACTGGGAGAACAGCTCCACATTCACGCCAGGTGAAGCTGAGCCCTGAGGCCAGGGAC	1340
3180	QY	CTTATTATTAACTTTGCGGAGGACCCGAGATCGCTTAGGCAAGATGTGCTGATGA	3239
1341	Db	CTCATCCACAGCTGTGCTCGCAGACACCAGCCCTGGGGCGGAATGGGCGCATGAC	1400
3240	QY	ATPAAAGCTCATCCATTTTTTAAACAAATTGACTTCTCCAGTCACTTGACACAGCAGTCT	3299
1401	Db	CTGAAGGCCACCCCTTCTCAGCGCCATTGACTTCTCCAGTGACATCCGGAAGCATCCA	1460
3300	QY	GCTTCATACATCTCTAAATCACACACCCACACAGATACATCAATTTTGNCTGTGTGAT	3359
1461	Db	GCCCTCAGTGTCCCATCCATCAGCACCCCATGGACACCTCGAATTCGACCCCGTAGAT	1520
3360	QY	CCTGATAAATATGAGTGTGATGAAGCAGGAAGAAATGTAATGACACTCTCAATGGA	3419
1521	Db	GAAGAAAGCCCTTGAAGGATGCCAGC--GAAGGTAGCACCAAGGCTGGGACACATC	1577
3420	QY	TGGTATAAATGGAAGCATCTGAAACATGTCATCTATGAATTTACCTTCGGAAGTTTT	3479
1578	Db	ACCTCGCCCAATAACAAGCATCTTGAGCAGCGCATTTTACGAATTCACCTTCGGAAGTTTC	1637
3480	QY	TTTGTATGACATGGCTACCCCATATAATTATCCGAAGCTTATTCGAATATGAATACATTAAT	3539
1638	Db	TTTGTATGACATGGCTACCCCTTTCGATGCCCAAGCCCTTCAGAGCAGAAAGCTTCACAG	1697
3540	QY	TCACAAAGGCTCAGA	3553
1698	Db	GCTGAGAGCTCAGA	1711

Search completed: January 16, 2003, 10:32:01
Job time : 562.163 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:02 : Search time 81.0582 Seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: 1 accttgggtgctgggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUTS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3984	100.0	3984	US-09-442-100-3	Sequence 3, Appli
2	2483	62.3	3213	US-09-442-100-5	Sequence 5, Appli
3	760	19.1	3155	US-09-442-100-7	Sequence 7, Appli
4	742.8	18.6	1961	US-09-509-902A-15	Sequence 15, Appli
5	675.8	17.0	1498	US-09-509-902A-6	Sequence 6, Appli
6	609.8	15.3	638	US-09-328-111-26	Sequence 26, Appli
7	469	11.8	5720	US-09-442-100-1	Sequence 1, Appli
8	452.8	11.4	678	US-09-328-111-66	Sequence 66, Appli
9	269.4	6.8	1935	US-08-878-989-11	Sequence 11, Appli
10	269.4	6.8	1935	US-09-272-796-11	Sequence 11, Appli
11	254.6	6.4	3018	US-08-860-150-6	Sequence 6, Appli
12	254.6	6.4	3018	US-09-338-132-6	Sequence 6, Appli
13	173.2	4.3	2160	US-09-588-256-1	Sequence 1, Appli
14	172.4	4.3	2101	US-08-860-150-1	Sequence 1, Appli
15	172.4	4.3	2101	US-09-338-132-1	Sequence 1, Appli
16	136.2	3.4	4739	US-08-685-871-1	Sequence 1, Appli
17	122.8	3.1	2706	US-08-630-822A-61	Sequence 61, Appli
18	122.8	3.1	2706	US-09-005-089-61	Sequence 61, Appli
19	122.8	3.1	2706	US-09-171-156A-20	Sequence 20, Appli
20	121.8	3.1	5053	US-08-685-576-2	Sequence 2, Appli
21	117	2.9	4363	US-08-685-576-5	Sequence 5, Appli
22	111	2.8	2311	US-08-712-709-6	Sequence 6, Appli
23	111	2.8	2311	US-09-111-444-6	Sequence 6, Appli
24	111	2.8	2311	US-09-541-228-6	Sequence 6, Appli
25	111	2.8	2370	US-09-031-295-1	Sequence 1, Appli
26	104	2.6	1599	US-09-256-465-1	Sequence 1, Appli
27	104	2.6	1599	US-09-167-322-3	Sequence 3, Appli

Sequence 1, Appli
Sequence 3, Appli
Sequence 48, Appli
Sequence 24, Appli
Sequence 2, Appli
Sequence 10, Appli
Sequence 1, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 12, Appli
Sequence 11, Appli
Sequence 11, Appli
Sequence 1, Appli
Sequence 25, Appli
Sequence 24, Appli
Sequence 25, Appli
Sequence 25, Appli

28 100.4 2.5 2196 1 US-08-313-274-1
29 92.6 2.3 2549 4 US-09-467-082-3
30 85.4 2.1 2244 3 US-09-094-714A-48
31 85.4 2.1 2245 4 US-09-223-749-24
32 76.6 1.9 1890 3 US-09-289-466-2
33 76.6 1.9 1891 3 US-09-016-000-10
34 75 1.9 1891 3 US-09-289-466-1
35 74.6 1.9 1732 4 US-09-430-564-1
36 74.6 1.9 2511 2 US-08-422-699A-8
37 74.6 1.9 2511 2 US-08-422-706B-8
38 74.6 1.9 2726 2 US-08-422-699A-12
39 74.6 1.9 2726 2 US-08-422-706B-12
40 74.6 1.9 3182 1 US-08-484-044-11
41 74 1.9 2104 4 US-09-313-930-1
42 72.8 1.8 1276 1 US-07-688-352C-25
43 72.8 1.8 1276 5 PCT-US91-02714-24
44 72.6 1.8 1273 2 US-08-474-379C-25
45 72.6 1.8 1273 3 US-09-146-249A-25

ALIGNMENTS

RESULT 1
US-09-442-100-3
: Sequence 3, Application US/09442100
: Patent No. 6359193
: GENERAL INFORMATION:
: APPLICANT: Xu, Tian
: APPLICANT: Tao, Wufan
: APPLICANT: Wang, Weiye
: APPLICANT: Zhang, Sheng
: APPLICANT: Yu, Wan
: TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/442,100
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/411,111
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Mistock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 6523-003
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3984 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: unknown
: MOLECULE TYPE: CDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 231..3623

US-09-442-100-3

Query Match	100.0%	Score 3984	DB 4	Length 3984
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3984	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	ACCTTTGGTGTCTGGACGGACTCTGGCCGCTCAGGCTCGCGCTCAGGCCGCTCAGGCCGCTGACC	60		
DB 1	ACCTTTGGTGTCTGGACGGACTCTGGCCGCTCAGGCTCGCGCTCAGGCCGCTCAGGCCGCTGACC	60		
QY 61	GCTGTCAGGAGCTCTGCTCTCCCTCCAGAGTAAATATATATATGTAAGAATTTTA	120		
DB 61	GCTGTCAGGAGCTCTGCTCTCCCTCCAGAGTAAATATATATATGTAAGAATTTTA	120		
QY 121	ACAGTCTGGGGACTTCTTGAAGGATCAATTTTCATCTTTGCTCAGAGAAGCTCTGGA	180		
DB 121	ACAGTCTGGGGACTTCTTGAAGGATCAATTTTCATCTTTGCTCAGAGAAGCTCTGGA	180		
QY 181	TCTATCAATTAAGAAGTCTCTTCTGTGGGTACATATATAGATGTTTTCATGAAGAGGA	240		
DB 181	TCTATCAATTAAGAAGTCTCTTCTGTGGGTACATATATAGATGTTTTCATGAAGAGGA	240		
QY 241	GTGAAAGCCAGAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAACATATA	300		
DB 241	GTGAAAGCCAGAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAACATATA	300		
QY 301	CTGTCAGTACGCGCAATGTTACAAGAAATTCGGGAATCCCTTAGGAAATTTATCTAAAC	360		
DB 301	CTGTCAGTACGCGCAATGTTACAAGAAATTCGGGAATCCCTTAGGAAATTTATCTAAAC	360		
QY 361	CATCTGATGCTGTAAAGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC	420		
DB 361	CATCTGATGCTGTAAAGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC	420		
QY 421	AAGTCAGAAATCCACCAGATTTGGGAGCATATAAAGCCTTCGAGGAATTCGAACACT	480		
DB 421	AAGTCAGAAATCCACCAGATTTGGGAGCATATAAAGCCTTCGAGGAATTCGAACACT	480		
QY 481	CTCTGCTTCCATTTGCAAAATGAACAAATCTTCTCGGAGTACTTCAGAACTTAATCCAC	540		
DB 481	CTCTGCTTCCATTTGCAAAATGAACAAATCTTCTCGGAGTACTTCAGAACTTAATCCAC	540		
QY 541	AAATGCTTCAAGACTTCCAGCTGCTGGATTTGATAGGATATGGTATACAAGCTCTTC	600		
DB 541	AAATGCTTCAAGACTTCCAGCTGCTGGATTTGATAGGATATGGTATACAAGCTCTTC	600		
QY 601	AGAAACTCAACAGAGATATAGAAGCAATTTGAATTCATTAGTAAATGAGTTACC	660		
DB 601	AGAAACTCAACAGAGATATAGAAGCAATTTGAATTCATTAGTAAATGAGTTACC	660		
QY 661	AAGATCTTCAGAGAGCAGATGGCTGCGAGCTGCCAGACCTATTAAATGCCAGCATGA	720		
DB 661	AAGATCTTCAGAGAGCAGATGGCTGCGAGCTGCCAGACCTATTAAATGCCAGCATGA	720		
QY 721	AACGAGGATATGTCAGCAATCAGTTAACGCCAAGAGCTGGAAGGTTCTAAGAAT	780		
DB 721	AACGAGGATATGTCAGCAATCAGTTAACGCCAAGAGCTGGAAGGTTCTAAGAAT	780		
QY 781	CTTATGTTCTCAGAGCATGGCCGCTCATTAGGAGAAAGTGGGCTATCATTTCTGAGA	840		
DB 781	CTTATGTTCTCAGAGCATGGCCGCTCATTAGGAGAAAGTGGGCTATCATTTCTGAGA	840		
QY 841	GTCCCAACTCACAGACAGATGTAGGAAGACTTTGCTGGATCTGGTATATCAGCATTTG	900		
DB 841	GTCCCAACTCACAGACAGATGTAGGAAGACTTTGCTGGATCTGGTATATCAGCATTTG	900		
QY 901	TTCAAGCTCACTTAGCAACGGACAGAGTGAACCCCCCACCACCACTCAAGTAAGGA	960		
DB 901	TTCAAGCTCACTTAGCAACGGACAGAGTGAACCCCCCACCACCACTCAAGTAAGGA	960		
QY 961	GTGTTACTCTCCACCACCTCCAGAGGCCAGACTCTCCCTTCCAAGAGGTACAACTCCAC	1020		
DB 961	GTGTTACTCTCCACCACCTCCAGAGGCCAGACTCTCCCTTCCAAGAGGTACAACTCCAC	1020		

Qy	1021	CTCCCCCTTCATGGGACCAAACTCTCAAAACAAAGCGTATTCTGGAACATGGAATACG	1080
Db	1021	CTCCCCCTTCATGGGACCAAACTCTCAAAACAAAGCGTATTCTGGAACATGGAATACG	1080
Qy	1081	TAAATCTCCGAAATCTCTCTGTGCCACCTGGGCGATGGCAAGAGGGCTATCCTCCACCAC	1140
Db	1081	TAAATCTCCGAAATCTCTCTGTGCCACCTGGGCGATGGCAAGAGGGCTATCCTCCACCAC	1140
Qy	1141	CTCTCAACACTTCCCCCATGAATCCCTTAATCAAGGACAGAGAGGCATTAGTCTGTTC	1200
Db	1141	CTCTCAACACTTCCCCCATGAATCCCTTAATCAAGGACAGAGAGGCATTAGTCTGTTC	1200
Qy	1201	CTGTTGGCAGAACCAATCATCATGCGAGTCTAGCAAAATTTAACTTTCCATCAGGGA	1260
Db	1201	CTGTTGGCAGAACCAATCATCATGCGAGTCTAGCAAAATTTAACTTTCCATCAGGGA	1260
Qy	1261	GACCTGGAATGCAGATGCTACTGGACAACCTGATTTCAATGATACACCAAAATTTGTGCC	1320
Db	1261	GACCTGGAATGCAGATGCTACTGGACAACCTGATTTCAATGATACACCAAAATTTGTGCC	1320
Qy	1321	CTGCTGGCACTGTGAATCGCGACGCCACCCTCCATATCTCTTGACAGCAGCTAATGSAC	1380
Db	1321	CTGCTGGCACTGTGAATCGCGACGCCACCCTCCATATCTCTTGACAGCAGCTAATGSAC	1380
Qy	1381	AAAGCCCTTCCTGCTTTTACAAACAGGGGGATCTGCTGCTCTTCGTCATATACAATGGAA	1440
Db	1381	AAAGCCCTTCCTGCTTTTACAAACAGGGGGATCTGCTGCTCTTCGTCATATACAATGGAA	1440
Qy	1441	GTATTCTCAGTCTATGATGGTGCCAAAACAGAAATAGTATCAATACATGGAACATATAACA	1500
Db	1441	GTATTCTCAGTCTATGATGGTGCCAAAACAGAAATAGTATCAATACATGGAACATATAACA	1500
Qy	1501	TTAGTGTACCTGGACTGCAAAACAAATTTGGCTTCAGTCTATCTTCCTCCAGCCCACTAT	1560
Db	1501	TTAGTGTACCTGGACTGCAAAACAAATTTGGCTTCAGTCTATCTTCCTCCAGCCCACTAT	1560
Qy	1561	CCCCGAGCAGTGGGCATGAAATCCCTTACATGCCAACCTTAACATACAGTGAAGTCAAAAT	1620
Db	1561	CCCCGAGCAGTGGGCATGAAATCCCTTACATGCCAACCTTAACATACAGTGAAGTCAAAAT	1620
Qy	1621	CTTTTAAATAACCAATTAGGAATAGAGCAAGTCACTCTGCTAAATTCACAGCCTTCGCTA	1680
Db	1621	CTTTTAAATAACCAATTAGGAATAGAGCAAGTCACTCTGCTAAATTCACAGCCTTCGCTA	1680
Qy	1681	CAACAGTCACTGCAATTTACACAGCTCCCTATTCAACAGCCTGTGAAAAGTATGCGTGAT	1740
Db	1681	CAACAGTCACTGCAATTTACACAGCTCCCTATTCAACAGCCTGTGAAAAGTATGCGTGAT	1740
Qy	1741	TAAACACAGAGCTACAGACTGCTTTTAGCACCTTACACACCCCTTCCTTGGATACCAAGCCAA	1800
Db	1741	TAAACACAGAGCTACAGACTGCTTTTAGCACCTTACACACCCCTTCCTTGGATACCAAGCCAA	1800
Qy	1801	TTCAAACTGTTCACCCAGTCCCTTTCTGAGGGAACCGCTTCAATGTGACTGTGATGC	1860
Db	1801	TTCAAACTGTTCACCCAGTCCCTTTCTGAGGGAACCGCTTCAATGTGACTGTGATGC	1860
Qy	1861	CACCTGTGCTGAAGCTCCAAACTATCAAGGACCACCAACCCCTTACCCAAACATCTGC	1920
Db	1861	CACCTGTGCTGAAGCTCCAAACTATCAAGGACCACCAACCCCTTACCCAAACATCTGC	1920
Qy	1921	TGCACCAAAACCCATCTGTTCTCCATACGAGTCAATCAGTAAAGCTTAGCAAAAGAGATC	1980
Db	1921	TGCACCAAAACCCATCTGTTCTCCATACGAGTCAATCAGTAAAGCTTAGCAAAAGAGATC	1980
Qy	1981	AGCCAAGCTTGCCCAAGGAAGTGAAGTGAAGAGATTATGAAAATGTTGATAGTGGGG	2040
Db	1981	AGCCAAGCTTGCCCAAGGAAGTGAAGTGAAGAGATTATGAAAATGTTGATAGTGGGG	2040
Qy	2041	ATAAGAAAAGAAACAGATTACAACTTCACCTATTACTGTTAGGAAAACAAAGAAAGATG	2100
Db	2041	ATAAGAAAAGAAACAGATTACAACTTCACCTATTACTGTTAGGAAAACAAAGAAAGATG	2100

QY 2101 AAGAGCGAAGGGAATCTCGTATTCAAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG 2160
 DB 2101 AAGAGCGAAGGGAATCTCGTATTCAAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG 2160
 QY 2161 AGCAACATGTAGAAATGTACTCAATCTCATCAGAGCGTCTACATCGTAAAAACAAT 2220
 DB 2161 AGCAACATGTAGAAATGTACTCAATCTCATCAGAGCGTCTACATCGTAAAAACAAT 2220
 QY 2221 TAGAGAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGA 2280
 DB 2221 TAGAGAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGA 2280
 QY 2281 TGCTTTGCCAAAAGAAATCTAATTACATCGTCTTAAAAGGGCTAAAAATGGACAAGTCTA 2340
 DB 2281 TGCTTTGCCAAAAGAAATCTAATTACATCGTCTTAAAAGGGCTAAAAATGGACAAGTCTA 2340
 QY 2341 TGTTTGTGAAGATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGTCTAGCAAGAA 2400
 DB 2341 TGTTTGTGAAGATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGTCTAGCAAGAA 2400
 QY 2401 AAGTAGATACTAAGGCTTTGTATGCAACAAAACTCTCCAAAAGAAAGATGTTCTTCTTC 2460
 DB 2401 AAGTAGATACTAAGGCTTTGTATGCAACAAAACTCTCCAAAAGAAAGATGTTCTTCTTC 2460
 QY 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
 DB 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
 QY 2521 GGGTAGTCTGCTATATTAATTCATCCAGATAAGGACAATTTATCTTTGTAATGAGCT 2580
 DB 2521 GGGTAGTCTGCTATATTAATTCATCCAGATAAGGACAATTTATCTTTGTAATGAGCT 2580
 QY 2581 ACATTCTGGGGTGATGATGATGAGCTTAAATTAAGATGGCACTTTCCAGAAAGTC 2640
 DB 2581 ACATTCTGGGGTGATGATGATGAGCTTAAATTAAGATGGCACTTTCCAGAAAGTC 2640
 QY 2641 TGCCACGATCTACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTT 2700
 DB 2641 TGCCACGATCTACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTT 2700
 QY 2701 TTATTCATAGAGATATAAACCCTGATAATTTTGTGATGCTGATGGTCATATTAAT 2760
 DB 2701 TTATTCATAGAGATATAAACCCTGATAATTTTGTGATGCTGATGGTCATATTAAT 2760
 QY 2761 TGACTGACTTTGGCCTCTGCACCTGCTCAGATGGACACACGATTTCTAAGTACTATCAGA 2820
 DB 2761 TGACTGACTTTGGCCTCTGCACCTGCTCAGATGGACACACGATTTCTAAGTACTATCAGA 2820
 QY 2821 GTGGTGACCATCCACGGCAAGATAGCATGGATTTCAATGAATGGGGGATCCCTCAA 2880
 DB 2821 GTGGTGACCATCCACGGCAAGATAGCATGGATTTCAATGAATGGGGGATCCCTCAA 2880
 QY 2881 GCTGTCGATGTGAGACAGACTGAAGCCATAGAGGGAGAGCTGCACGCCAGCAGC 2940
 DB 2881 GCTGTCGATGTGAGACAGACTGAAGCCATAGAGGGAGAGCTGCACGCCAGCAGC 2940
 QY 2941 GATGTCATAGCACATCTTTGGTGGGACTCCCAATATATTTGCACCTGAAGTGTGCTAC 3000
 DB 2941 GATGTCATAGCACATCTTTGGTGGGACTCCCAATATATTTGCACCTGAAGTGTGCTAC 3000
 QY 3001 GAACAGGATACACAGAGTTGTGTGATTTGGTGGAGTGTGTTATCTTTTGAATGT 3060
 DB 3001 GAACAGGATACACAGAGTTGTGTGATTTGGTGGAGTGTGTTATCTTTTGAATGT 3060
 QY 3061 TGGTGGGACAACTCTTTCTGGACAAACACACCATTTAGAAACACAAATGAAGTTATCA 3120
 DB 3061 TGGTGGGACAACTCTTTCTGGACAAACACACCATTTAGAAACACAAATGAAGTTATCA 3120
 QY 3121 ACTGGCAAAACATCTTTCATCTCCACCAAGCTAACTCACTGCTGAAGCTTCTGATC 3180
 DB 3121 ACTGGCAAAACATCTTTCATCTCCACCAAGCTAACTCACTGCTGAAGCTTCTGATC 3180
 QY 3181 TTATTATTAACTTTGCCGAGGACCGGAAGATCGCTTAGCAAGAATGGTCTGATGAAA 3240

DB 3181 TTATTATTAACTTTGCCGAGGACCGCAAGATCGTTAGCAAGAATGGTCTGATGAAA 3240
 QY 3241 TAAAAGCTCATCTCAATTTTAAACAATTTGACTTCTCCAGTGACCTGAGACAGAGTCTG 3300
 DB 3241 TAAAAGCTCATCTCAATTTTAAACAATTTGACTTCTCCAGTGACCTGAGACAGAGTCTG 3300
 QY 3301 CTTCATACATCTCTTAAATCACACACCACAGATACATCAAAATTTTGTCTGTTGATC 3360
 DB 3301 CTTCATACATCTCTTAAATCACACACCACAGATACATCAAAATTTTGTCTGTTGATC 3360
 QY 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAATGACACTCTCAATGGAT 3420
 DB 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAATGACACTCTCAATGGAT 3420
 QY 3421 GGTATAAATGGAAGCATCTCGAATGATGATCAATCTTCTATGAATTTACCTCCGAGGTTTT 3480
 DB 3421 GGTATAAATGGAAGCATCTCGAATGATGATCAATCTTCTATGAATTTACCTCCGAGGTTTT 3480
 QY 3481 TTGATGACATGCTACCCATATAATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3540
 DB 3481 TTGATGACATGCTACCCATATAATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3540
 QY 3541 CACAAGGCTCAGAGCAGCAGTCCGATGAAGATGATCAAAAACACAGAGCTCAGAGATTTAAA 3600
 DB 3541 CACAAGGCTCAGAGCAGCAGTCCGATGAAGATGATCAAAAACACAGAGCTCAGAGATTTAAA 3600
 QY 3601 ATCCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGATTTGTAAG 3660
 DB 3601 ATCCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGATTTGTAAG 3660
 QY 3661 GGCTGAAATGCCAGGTTTTCAGGTTCTGAGAGTAAATTTATGCAAAATATGACAGC 3720
 DB 3661 GGCTGAAATGCCAGGTTTTCAGGTTCTGAGAGTAAATTTATGCAAAATATGACAGC 3720
 QY 3721 TATATATGCTGCTCTGTGTACAAATTTTATTTTCTTAAATTTATGGAATCCTTTTAA 3780
 DB 3721 TATATATGCTGCTCTGTGTACAAATTTTATTTTCTTAAATTTATGGAATCCTTTTAA 3780
 QY 3781 AATGTTAATTTATCCAGCGTTTAAATCAGTATTAGAAAAAATGTTTAAAGGAAG 3840
 DB 3781 AATGTTAATTTATCCAGCGTTTAAATCAGTATTAGAAAAAATGTTTAAAGGAAG 3840
 QY 3841 TAAATATGAACCTGAATATTATAGTCTGTTGTTTAAAGTACTTAAATTAAGTAG 3900
 DB 3841 TAAATATGAACCTGAATATTATAGTCTGTTGTTTAAAGTACTTAAATTAAGTAG 3900
 QY 3901 TGCTTTGTTTAAAGGAGAAACCTGTTATCTATTGTTATATGCTAAATTTTAAAA 3960
 DB 3901 TGCTTTGTTTAAAGGAGAAACCTGTTATCTATTGTTATATGCTAAATTTTAAAA 3960
 QY 3961 TACAAGAGTTTTTGAATTTTTT 3984
 DB 3961 TACAAGAGTTTTTGAATTTTTT 3984

RESULT-2

US-09-442-100-5

; Sequence 5, Application US/09442100

; Patent No. 6359193

; GENERAL INFORMATION:

; APPLICANT: Xu, Tian

; APPLICANT: Tao, Wufan

; APPLICANT: Wang, Welyi

; APPLICANT: Zhang, Sheng

; APPLICANT: Yu, Wan

; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

Fri Jan 17 11:16:54 2003

STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0; Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100

Query Match 62.3%; Score 2483; DB 4; Length 3213;
Best Local Similarity 87.4%; Pred. NO. 0;
Matches 2750; Conservative 0; Mismatches 385; Indels 15; Gaps 4;

Qy	732	GTGCGAATCAGTTAACCGCAACAGAGCTTGGAAAGGTTCTTAAGAAATCCTTAGTTCTCT	791
Db	1	GTGCAACATTCATTAATCAACGAACAAAGCTTGGAAAGGTTCTTAAGAGTCTCTAGTTCTCT	60
Qy	792	CACAGGATGGCGCCGACCTAGGAGAAAAGTTGGCCCTATCATCTGAGAGTCCCAACTCA	851
Db	61	CACGACACGGCCCATCTCTAGGAGAAAATGTGTTTATCGTTCTGAAAGCCCAACTCA	120
Qy	852	CACAGAGATGTAGGAAGACCTTTGTTCTGGATCTGGTATATCAGCATTTGTTCAAGTTCAC	911
Db	121	CAGCGGATGTAGGAAGACCTCTGTCTGTATCCGCATTCGACATTTGCTCAAGTTCAC	180
Qy	912	CCTAGCAACGACAGAGAGTGAACCCGCCACACACCTCAAGTAAGAGTGTTACTTCCT	971
Db	181	CCRAGCAATGGACAGAGAGTGAACCCGCCACACACCTCAAGTTAGAGTGTTACTTCCT	240
Qy	972	CCACCACCTCCAGAGGCCAGACTCCCTCCCAAGAGTACAACCTCCACACTCCCCCTTCA	1031
Db	241	CCACCACCTCCGAGAGGCCAGACCCACCTCCCGAGGCACCACTCCCTCCCTCCCTCA	300
Qy	1032	TGGGAACCAACTCTCAACAAAGCGCTATCTGGAACATAGGAATAGTAACTCCCGA	1091
Db	301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGACATGGAGTAGTAACTCCCGA	360
Qy	1092	ATCTCTCCTGCCACCTGGGCGATGGCAAGAGGCTATCCTCCACACACCTCTCAACACT	1151
Db	361	ATCTCCCTCTGTTCACCTGGGCGTGGCAGAGGGGTACCTCCACACACTCTTACCAC	420
Qy	1152	TCCCCCATGAATCCTCCTAATCAAGGACAGAGAGGCATTAGTTCTGTCTGTTGGCAGA	1211
Db	421	TCTCCCATGAATCCCTCCTACCCAGGCTCAGAGGGCCATTAGTCTGTGTTCCAGTGGTAGA	480
Qy	1212	CAACCAATCATGCAGAGTTCTACCAAAATTTAACTTTCCATCAGGAGACCTTGGGAATG	1271

QY 2352 ATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGATACT 2411
DB 1618 ATAAAGACATAGGAATAGGAGCGTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGATACT 1677
QY 2412 AAGGCTTTGTATGCAACAAAACACTCTTCGAAAGAAAGATGTTCTCTCGAAATCAAGTTC 2471
DB 1678 AAGGCTTTGTATGCAACAAAACACTCTTCGAAAGAAAGATGTTCTCTCGAAATCAAGTTC 1737
QY 2472 GCTCATGTTAAAGCTGAGAGATATCTCGGCTGAAGCTGACAATGAATGGGTAGTTCCT 2531
DB 1738 GCTCATGTTAAAGCTGAGAGATATCTCGGCTGAAGCTGACAATGAATGGGTAGTTCCT 1797
QY 2532 CTATATATTTTCAAGATAAGGACAATTTATCTTTTGAATGAGCTACATTTCTCTGGG 2591
DB 1798 CTGTACTCTCTTCCAGCACAAGGACAATTTGATCTTTGTGATGAGCTACATTTCTCTGG 1857
QY 2592 GGTGATATGATGAGCTTATTAATAGATGGGATCTTTCCAGAAAGTCTGGCAGATTC 2651
DB 1858 GGGATATGATGAGCTTATTAATAGATGGGATCTTTCCAGAAAGTCTGGCAGATTC 1917
QY 2652 TACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTTTATTCATAGA 2711
DB 1918 TACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTTTATTCATAGA 1977
QY 2712 GATATTAACCTGATAATATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2771
DB 1978 GATATTAACCTGATAATATTTTGTGATGATGATGATGATGATGATGATGATGATGAT 2037
QY 2772 GGCCTCTGACCTGGCTTCAGATGGACACACAGATCTTAAGTACTATCAGAGTGGTACCAT 2831
DB 2038 GGCCTCTGACCTGGCTTCAGATGGACACACAGATCTTAAGTACTATCAGAGTGGTACCAT 2097
QY 2832 CCACGCAAGATAGATGATTTTCAATGATGATGATGATGATGATGATGATGATGATGAT 2891
DB 2098 CCACGCAAGATAGATGATTTTCAATGATGATGATGATGATGATGATGATGATGATGAT 2157
QY 2892 GGAGCAGACTGAAGCATTAGAGCGGAGAGCTGCACGCGCAGCAGCAGCTGCTTAGCA 2951
DB 2158 GGGCAGACTGAAGCATTAGAGCGGAGAGCTGCACGCGCAGCAGCAGCTGCTTAGCC 2217
QY 2952 CATCTTTGGTGGACTCCCAATATATTTGACCTGAAGTGTGCTAGCAGCAGATAC 3011
DB 2218 CATCTTTGGTGGACTCCCAATATATTTGACCTGAAGTGTGCTAGCAGCAGATAC 2277
QY 3012 ACACAGTCTGTGATGGTGGAGTGTGCTGATTTTCAATGATGATGATGATGATGATGAT 3071
DB 2278 ACACAGTCTGTGATGGTGGAGTGTGCTGATTTTCAATGATGATGATGATGATGATGAT 2337
QY 3072 CCTCCTTTCTGGCACAACACCATTTAGAAACACAAATGAAGTTATCAACTGGCAACA 3131
DB 2338 CCTCCTTTCTGGCACAACACCATTTAGAAACACAAATGAAGTTATCACTGGCAACT 2397
QY 3132 TCTCTTACATTTCCACCAACGCTAACTCAGTCTGAGAGCTTCTGATCTTATTAATAA 3191
DB 2398 TCTCTACACATCTCTCCTCAAGCTAAGTGTGATGATGATGATGATGATGATGATGATGAT 2457
QY 3192 CTTTCCGAGGAGCAGCAGATGCTTAGCAAGATGGTGTGATGATGATGATGATGATGATGAT 3251
DB 2458 CTGTGCGAGGAGCAGCAGCCTCGGCAAGAGAGGCTGATGATGATGATGATGATGATGAT 2517
QY 3252 CCATTTTAAACCAATTTGACTTCTCCAGTGCAGTGCAGCAGCTGCTTCTCATACAT 3311
DB 2518 CCATTTTAAACCAATTTGACTTCTCTAGTGCAGTGCAGCAGCTGCTTCTCATACAT 2577
QY 3312 CTTAAATCAGACACCCCAACAGATACATCAAAATTTGATCTGTTGATCTCTGATAAATTA 3371
DB 2578 CTTAAATCAGCAGTCCCAACAGATACATCAAAATTTGATCTGTTGATCTCTGATAAATTA 2637
QY 3372 TGGAGTATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3431
DB 2638 TGGAGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2697

QY 3432 GGAAAGCATCCCTGAACATGCTATGATTTACCTCCGAGGTTTTTTTGTATGACAAAT 3491
DB 2698 GGGAACACCCCGAGCAGCTTTCTATGATGATGATGATGATGATGATGATGATGATGAT 2757
QY 3492 GGTACCCCATATAATATCCGAGCCTATGATGATGATGATGATGATGATGATGATGATGAT 3551
DB 2758 GGTACCCCATATAATATCCGAGCCTATGATGATGATGATGATGATGATGATGATGATGAT 2817
QY 3552 GAGCAGCAGCTCGAGTGAAGATGATCAACACAGAGCTGATGATGATGATGATGATGAT 3611
DB 2818 GAACAACAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877
QY 3612 GTATATGTTTAAACACTAGTAAATGATGATGATGATGATGATGATGATGATGATGATGAT 3671
DB 2878 GTGATGTTTAAATAGGAGTCAATGTAAGTAAATGATGATGATGATGATGATGATGATGAT 2933
QY 3672 CGAGGTGTTTTCAGGCTGAGAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 3731
DB 2934 CAGGGGTTTTTGAAGTTTTGAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 2985
QY 3732 GCTCTGTGATCAATATTTTATTTTCTTAATGATGATGATGATGATGATGATGATGATGAT 3791
DB 2986 GCTCTGTGATCAATATTTTATTTTCTTAATGATGATGATGATGATGATGATGATGATGAT 3045
QY 3792 ATTCCAGCGCTTAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3851
DB 3046 ATTCCAGCGCTTAAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3105
QY 3852 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3900
DB 3106 CTGATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3154

RESULT 3

US-09-442-100-7
Sequence 7, Application US/09442100
Patent No. 6359193

GENERAL INFORMATION:

APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Weli
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111

FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 3155 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2943

US-09-442-100-7

Query Match 19.1%; Score 760; DB 4; Length 3155;

Best Local Similarity 63.7%; Pred. No. 5.2e-188;

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

Qy 1653 CACTCTGCTAATCTCAGCCCTTCTGCTACACAGTCTACTGCAATATACACAGCTCTCTATT 1712
 Db 985 CAACCTGAGCCCTCACTGCGCGCCCAACACAGCTACCGCGCTGAGCGCGCCGACACATC 1044
 Qy 1713 CAACAGCCTGTGAAGATGCTGCTGATTAACACAGCTACAGACTGCTTTAGCACCT 1772
 Db 1045 CTTACCCCTGTGAAGAGCGTGTGCTGCGCGCCGAGCCCGAGACAGCCGCTGGGCGCC 1104
 Qy 1773 ACACACCTCTCTTGATACACAGCCCAATTCAAACTGTTCAAACCCAGTCTCTTTCTCTGAG 1832
 Db 1105 TCGACACCCGCTGGTGGCTGCGCCACACAGCCTGCCACTGAGCCCTGGAGACGAG 1164
 Qy 1833 GGAACCGCTTCAAAATGTGACTGTGATGCCACCTGTGCTGAAGCTCCAACTATCAAGGA 1892
 Db 1165 GAGGCGCAGCGCCACACACCGCTGTGACTATGCGGCTCCGAGCGCAGGTGC 1224
 Qy 1893 CCACACACCCCTACCCAAACATCTGCTGCACCAACACCATCTGTCTCTCCATACGAG 1952
 Db 1225 CCACCGCTCTGCTATCCAAAGCCTTGTGCTGCCAGTAAGTCTGACGACGTACAGCTG 1284
 Qy 1953 TCAATCAGTAAGCTAGCAAGAGAGATCAGCCAAAGCTTGCCTCAAGAGAGATGAGAGTAA 2012
 Db 1285 GACCTGGACGCTGTGCACCACTGTGCAGCAGAGTCTGCGAGGGGCGACTGATCTAGAC 1344
 Qy 2013 AAGAGTTA-----TGAAATGTTGATGTTGGGATGAAGATAAGAAAGAACAG 2057
 Db 1345 GGGAGTCAAGAGCGCACAAAGGTGCGAAGGGAGACAAAGTGGCAGAGACAAAGAGCAG 1404
 Qy 2058 ATTACAACCTTCACCTATTACTGTTAGGAAACAAAGAAAGATGAAGCGGAGGAATCT 2117
 Db 1405 ATTACAGACTCCCGGCTGCTGCTCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 Qy 2118 CGTATTCAAGTTATTCCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAT 2177
 Db 1465 CGCATCAAGAGTTACTCCCTTATCCCTTCAATTTCTTCATGGAGCAACAGCTGGAGAT 1524
 Qy 2178 GTACTCAATCTCATCAGCAGCGTCTACATCGTAAACAAATTAAGAAATGAATGATG 2237
 Db 1525 GTCATCAAAACCTACACGACGAAGTCTAGCGGAGGCTACAGCTGGAGCAGGAATGCC 1584
 Qy 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGGCAAAAGAA 2297
 Db 1585 AAAGCTGGCTCTGTGAGCGCGCAGCAGCAGATGAGGAAGTCTCTACCAAGAGGAG 1644
 Qy 2298 TCTAATACATCCGCTTTAAAGGGCTAAATGGACAGTCTATGTTTGTGAAGTAAG 2357
 Db 1645 TCTAATCAACCGCTGAAGAGGGCCCAAGATGGACAGTGCATGTTTGTGAAGATCAAG 1704
 Qy 2358 ACACATAGGATAGGACATTTGGTCAAGTCTGTCTAGCAAGAAATAGATACTAAGCT 2417
 Db 1705 ACTCTAGGCATCGGTCCTTTGGGGAAGTGTGCCCTGTGTAAGCTGGACACTCACGCT 1764
 Qy 2418 TTGATGCAACAAACATCTCTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTCGCTCAT 2477
 Db 1765 CTGTAGCCATCAAGACTCTCAGGAGAGAGATGCTCTGAACCGGAATCAAGTGGCCCAT 1824
 Qy 2478 GTTAAGGCTGAGAGAGATATCCTGCTGAAGCTGACATGAATGAATGGTGTCTCTATAT 2537

Db 1825 GTCAAGGCTGAGAGGACATCTCTGGCTGAAGCAGACAATGAGTGGTGAACACTCTAC 1884
 Qy 2538 TATTCAATCAAGATAGGACAAATTTATCTTTGTAATGACTACATTTCTCTGGGGTGAT 2597
 Db 1885 TACTCTTTCCAGGACAGGACGCCCTGTACTTTGTGATGACTACATACCAGCGGGGAT 1944
 Qy 2598 ATGATGAGCCTATTAAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGATTTACATA 2657
 Db 1945 ATGATGAGCCTGCTGATCAGGATGAGGCTCTCCCTGAGCACCTGGCCGCTTCTACAT 2004
 Qy 2658 GCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGGTTTTTATTATAGATATT 2717
 Db 2005 GCAGAGTTGACCTGGCCATTGAAAGTGTCCACAAGATGGGCTTTATCCACCGGACATC 2064
 Qy 2718 AAACCTGATAAATTTTGTGATGCTGATGCTCATATTAATTAATGACTGACTTTGGCCTC 2777
 Db 2065 AAGCCTGACAACATCTATCAGCTGAGTGGTGTATTAAGCTGACAGATTTTGGCCTC 2124
 Qy 2778 TGCACCTGGCTTCAGATGGACACACAGATTTCTAAGTACTATCAGAGTGGTACCATCCAGG 2837
 Db 2125 TGCACCTGATTCAGTGGACTCAATTTCCAAGTACTACCAAGAAAGGAAACCATGAGA 2184
 Qy 2838 CAAGATAGCATGATTTCAAGTAAATGAGGGGATCCCTCAAGCTGCTGATGGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCCGCTGACCTCTGGACGATGTTTCCAACCTGTGCTGTGGAGAC 2244
 Qy 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCACAGGAGTGTCTAGACATCTCT 2957
 Db 2245 AGTTAAAGACCTTGGAGCAGAGAGGCGCCAGACAGCACAGGAGTCTTGGCACATCTCT 2304
 Qy 2958 TTGTTGGGACTCCCAATTTATTTGACCTTGAAGTGTGTTAGCAAGAGGATACACAG 3017
 Db 2305 CTGTGCGGACACCAATTTACATCGCTCCGAGGTCTTCTCCGCAAGAGGTACACAG 2364
 Qy 3018 TTGTGCTATTGGTGGAGTGTGTTTATTTCTTTTGAATGTTGGTGGGACACCTCTCT 3077
 Db 2365 CTCTGTGACTGGTGGAGCGTGGTGTGATTTCTCTTTGAGATGCTGGTGGGACGCCCT 2424
 Qy 3078 TTCTTGGCAACCAACACCATTTAGAAACACAAATGAAGGTTATCAACTGGCAACATCTCT 3137
 Db 2425 TTCTTGGCCCCCACCACACAGACAGCAGCTGAAGCTGATCACTGGAGAGCACCTG 2484
 Qy 3138 CACATCCACACAGCTAACTCAGTCTCTGAGCTCTGATCTTATTTAATTTTGC 3197
 Db 2485 CATATCCCTACGCAAGTGAAGCTCAGGCTCAGGCTGAGCCCGAGACCTCATCAGAGCTGTC 2544
 Qy 3198 CGAGACCCGAGATCGCTTAGGCAAGAAATGCTGCTGATGAAATAAAGCTCATCTCCATTT 3257
 Db 2545 TGGCGGCTGACTCGCGCTGGGAGGATGGGAGATGAGCTCAAGGCACACCCGCTC 2604
 Qy 3258 TTTAAACAAATTTGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATCATCTTAA 3317
 Db 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGAGGCTGCACCTACGCTCCCAACC 2664
 Qy 3318 ATCACACACCAACAGATACATCAAAATTTGATCTCTGATCTCTGATAAATTTATGAGT 3377
 Db 2665 ATCAGCCACCCCTGGACACCTCCAAATTTTACCCGTTGAGATGAAGAAAGCCCTGGCAC 2724
 Qy 3378 GATGATACGAGGAGAAATGTAATGACACTCTCAATGATGATGATAAATAATGGAAG 3437
 Db 2725 GAGGCCAGC---GGAGAGGCGCAAGCCCTGGGACAGCTGGGCTCCCCCAGCAGCAAG 2781
 Qy 3438 CATCTTGAACATGATCTATGAATTTACCTTCCGAAGGTTTTTTTGTGATGACATGGCTAC 3497
 Db 2782 CATCCAGCAGCAGCTTCTATGAGTTCACCTTCCGAGGTTCTTCGATGACACAGGCTAT 2841
 Qy 3498 CCATATATTTATCCGAAGCC 3517
 Db 2842 CCTTCCGTTGCCGAAGCC 2861

RESULT 4

US-09-509-902A-15

; Sequence 15, Application US/09509902A

; Patent No. 6387676

; GENERAL INFORMATION:

; APPLICANT: Virca, Duke

; APPLICANT: Bird, Timothy A.

; APPLICANT: Anderson, Dirk M.

; APPLICANT: Marken, John S.

; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions

; FILE REFERENCE: 2877-US

; CURRENT APPLICATION NUMBER: US/09/509,902A

; CURRENT FILING DATE: 1999-08-03

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1961

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-509-902A-15

Query Match 18.6%; Score 742.8; DB 4; Length 1961;
Best Local Similarity 66.2%; Pred No. 1.2e-183;
Matches 1122; Conservative 0; Mismatches 557; Indels 15; Gaps 3;

QY 1872 GAAGCTCCAAACTATCAAGGACCAACACCCCTACCCAAAACATCTGCT-----GCACC 1926
DB 21 GGAGGCCAGACGGAGGTGCCCGCTCGCCCTACCCGAAGCACCTGCTGCTGGCAGC 80
QY 1927 AAACCCATCTGTTCTCCATAGAGTCAATCAGTAAGCCTAGCAAGAGGATCAG-----1982
DB 81 AAGTGGAGCAGTACGACCTGGACAGCCTGTGCGAGGATGGAGCAGAGCCTCCGTGGC 140
QY 1983 ---CCAAGCTGCCCAAGAGATGAGAGTGAAGAGAGTTATGAAGATGTTGATAGTGG 2039
DB 141 GSCCCAGACGCCGAGGCGGCGGCAAGAGCCGCAAGAGCCGAGGGGACAAAGCC 200
QY 2040 GATAAGAAAAGAAACAGATTACAACTTCCACTTATTCTTCTCAAGCATTTAAATCTTTATG 2099
DB 201 GGAAGGATAAAAGCAGATTTCAGACCTCTCCGTTCCCGTCCGCAAAAACAGCAGAC 260
QY 2100 GAAGCGGAAGGAATCTCGTATTCAAGTATTCTTCTCAAGCATTTAAATCTTTATG 2159
DB 261 GAAGAGACAGAGAGTACGATCAAGAGTCTACGACCTACTCGCCATAGCCCTTTAAGTCTTCATG 320
QY 2160 GAGCAACATGTAGAAATCTACTCAATCTCATCAGCAGCTCTACATCGTAAACAA 2219
DB 321 GAGCAGCAGCTGGAGATGATCAAACTTACAGCAAGATTAACCGAGGCTGCAG 380
QY 2220 TTAGAGATGAATGATGCGGGTTGGATTTATCTCAAGATGCCAGGATCAAAATGAGAAAG 2279
DB 381 CTGGAGCAAGAAATGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATGCGGAAG 440
QY 2280 ATGCTTTGCCAAAAGAAATCTAATTCATCCGTTTAAAGGGCTTAAATGACAGTCT 2339
DB 441 ATCTCTTACCAAGAGAGTCTAATTCACAGAGTTAAAGAGGGGCCAAGATGACAGTCT 500
QY 2340 ATGTTTGTCAAGATAAGACACTAGGAATAGGAGATTTGGTGAAGTCTGTAGCAAGA 2399
DB 501 ATGTTTGTCAAGATCAAAACCTGGGGATCGGTGCGCTTTGGAGAGTGTCCCTTGTCT 560
QY 2400 AAAGTAGACTAAGCTTTTGTATGCAACAAAACCTTTCGAAAAGAGATGTTCTTCTT 2459
DB 561 AAGGTGGACACTCACGCCCTGTACGCCATGAAGACCCCTAAGGAAAAAGATGTCTGTAAC 620
QY 2460 CGAATCAAGTCGTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAAATGA 2519
DB 621 CGGAATCAGGTGGCCACCTCAAGGCCGAGAGGGAGATCCTGGCCGAGGACAGCAATGAG 680
QY 2520 TGGGTAGTTCGCTATATTATTCATTCCAAGATAGGACAATTTACTTTTGTATGGAC 2579
DB 681 TGGGTGGTCAAACTCTTACTTCTTCCCAAGACAAGACAGGCTGTACTTTTGTGGAGC 740
QY 2580 TACATTCTTGGGGTGTATATGATGAGCCTTATTAATTAAGATGGCATCTTTCCAGAAAGT 2639

DB 741 TACATCCCTGGTGGGACATGATGAGCCCTGCTGATCCGATGGAGTCTTCCCTGAGCAC 800
QY 2640 CTGGCAGCAGATTCTACATAGCAGAACTTACTGTGCGATTGAAAGTGTTCATATAAATGGT 2699
DB 801 CTGGCCCGGTTCTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCACAGATGGC 860
QY 2700 TTATTTCATAGAGATATTAAACCTGATAATATTTTCATTGATGCTGATGCTCATATATAA 2759
DB 861 TTCATCCACCGAGACATCAAGCCTGATAACATTTTGTATAGATCTGGATGGTCCACATATA 920
QY 2760 TTGACTGACTTTTGGCCTCTGCACTTTCAGATGGACACACATTTCTTAAGTACTATCAG 2819
DB 921 CTCACAGATTTCCGCCCTCTGCACTTTCAGTGGTTCAGTGGACTCACAATTTCCAAATATTACCAG 980
QY 2820 AGTGGTGACATCCAGCGCAAGATAGCATGGATTTTCAGTAAATGAATGGGGGATCCCTCA 2879
DB 981 AAAGGAGCCATGTCAGACAGGACATGGAGCCGAGCCTCTGGATGATGTGTCT 1040
QY 2880 AGCTGTCGATGTGGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGACACAG 2939
DB 1041 AACTGTGCTGTGGGACAGAGCTGAAGACCCCTAGACAGAGGGCGGAGCAGCAGCACAG 1100
QY 2940 CGATGCTCTAGCACATCTTTTGGTGGGACTCCCAATATATTTCACCTGAAAGTGTTCATA 2999
DB 1101 AGTGTCTGGCACATCTACTGTGGGACTCCAAACTACATGCACCCGAGGTGCTCCTC 1160
QY 3000 CGAACAGGATACACACAGTTGTGATTTGGTGGAGTGTGGTGTATTCTTTTGAATG 3059
DB 1161 CGAAAGGATACACTCACTCTGTGACTGTGGAGTGTGGAGTGTATTCTTCGAGATG 1220
QY 3060 TTGGTGGGACAACTCTCTTTTGGCACAAACACCACTTAGAAACACAAATGAAGTGTATC 3119
DB 1221 CTGGTGGGACGCGCCCTTTTGGCACCTACTCCACAGAAACCCAGCTGAAGTGTATC 1280
QY 3120 AACTGCCAAACATCTTTCACATTCACCAAGCTAAACTCAGTCTGAAAGTCTTCTGAT 3179
DB 1281 AACTGGGAAACACAGCTCCACATTCAGCCAGGTGAAGTGAAGCTGAGCCTGAGCCAGGAC 1340
QY 3180 CTATTATTAACTTTGCCGAGGACCCGAAGATCGCTTAGGAGAGATGCTGCTGATGA 3239
DB 1341 CTCATCACCAAGCTGTGCTCTCCGACAGCACCCGCTGGGCGGAATGGGCGCATGAC 1400
QY 3240 ATAAAGCTCATCCATTTTAAACAAATGACTTCTCAGTGACCTGAGACAGCAGTCT 3299
DB 1401 CTGAAGGCCACCCCTTCTTCAGGCCATGACTTCTCAGTGACATCCGAGAGCATCCA 1460
QY 3300 GCTTCATACATTTCTTAAATTCACACACCCACAGATACATCAAAATTTTGTGTTGAT 3359
DB 1461 GCGCCCTAGCTTCCCAACCATCAGCCACCCCATGGACACCTCGAATTTGCACCCGCTAGAT 1520
QY 3360 CTTGATAAATTTGAGTGTATGATACAGGAGGAGAAATGTAAATCAGACTCTCAATGGA 3419
DB 1521 GAAGAGCCCTTGGAAAGCATCTGAACATGCTTATGAATTTACCTTCGGAAGTCT 1577
QY 3420 TGTATATAAATGGAAGCATCTGAACATGCTTATGAATTTACCTTCGGAAGTCT 3479
DB 1578 ACCTGCCCAATAACAGAGCATCTGAGCAGCATTTTACGAATTCACCTTCGGAAGTCT 1637
QY 3480 TTGTATGACAATGGCTTACCCATATAATTTCCGAAAGCTATTGAATATGAATACATTAAT 3539
DB 1638 TTGTATGACAATGGCTTACCCCTTTCATGATGCCAAAGCCCTTCAGGAGCAGAAAGTTCACAG 1697
QY 3540 TCACAGGCTCAGA 3553
DB 1698 GCTGAGAGCTCAGA 1711

RESULT 5:
US-09-509-902A-6
; Sequence 6, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 26

LENGTH: 638

TYPE: DNA

ORGANISM: Homo sapiens

US-09-328-111-26

Query Match

Best Local Similarity 98.4%; Score 609.8; DB 4; Length 638;

Matches 627; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 2178 G T A C T C A A A T C T C A T C A G C A G G C T T A C A T C G T A A A A A C A A T T A G A A T G A A T G A T G 2237
D b 637 G T A C T C A A A T C T C A T C A G C A G G C T T A C A T C G T A A A A A C A A T T A G A A T G A A T G A T G 578
QY 2238 C G G T T G G A T T A T C T C A A G A T C C C A G G A T C A A T C A A T C A A T G A A A G A T C T T T G C C A A A A G A A 2297
D b 577 C G G T T G G A T T A T C T C A A G A T C C C A G G A T C A A T C A A T G A A A G A T C T T T G C C A A A A G A A 518
QY 2298 T C T A A T T A C A T C C G T C T T A A A A G G C T A A A T G A C A A G T C T A T G T T T G T G A A G A T A A A G 2357
D b 517 T C T A A T T A C A T C C G T C T T A A A A G G C T A A A T G A C A A G T C T A T G T T T G T G A A G A T A A A G 458
QY 2358 A C A C T A G G A T A G G A C A T T T G T G A A G T C T G T C T A G C A A A A A A G T A G A T A C T A A G G C T 2417
D b 457 A C A C T A G G A T A G G A C A T T T G T G A A G T C T G T C T A G C A A A A A A G T A G A T A C T A A G G C T 398
QY 2418 T T G T A T G C A C A A A A A C T T C G A A A A A A G A T G T T C T T C G A A A T C A A T C A A G T C G C T C A T 2477
D b 397 T T G T A T G C A C A A A A A C T T C G A A A A A A G A T G T T C T T C G A A A T C A A T C A A G T C G C T C A T 338
QY 2478 G T T A A G G C T G A G A G A T A T C T G C T G A G C T G A C A A T G A A T G G T A G T T C G T C T A T A T 2537
D b 337 G T T A A G G C T G A G A G A T A T C T G C T G A G C T G A C A A T G A A T G G T A G T T C G T C T A T A T 278
QY 2538 T A T T A T C C A G A T A A G G A C A A T T A T A C T T T G T A A T G A C T A C A T T C C T G G G G T G A T 2597
D b 277 T A T T A T C C A G A T A A G G A C A A T T A T A C T T T G T A A T G A C T A C A T T C C T G G G G T G A T 218
QY 2598 A T G A T C A G C C T A T A A T T A G A T G G C A C T T T C C A A A A G T C G C A A G A T T C T A C A T A 2657
D b 217 A T G A T C A G C C T A T A A T T A G A T G G C A C T T T T C C A A A A G T C G C A A G A T T C T A C A T A 158
QY 2658 G C A G A A C T T A C C T G T C A G T T G A A A G T G G T T T C A T A A A A T G G T T T A T T C A T A G A T 2714
D b 157 G C A G A A C T T A C C T G T C A G T T G A A A G T G T C C T T A A A T G G T T T A T T C A T A G A T 98
QY 2715 A T T A A C C T G A T A A T T T T G A T T G A T G C G A T G G T C A T A T T A A A T G A C T A C T T T G G C 2774
D b 97 A T T A A C C T G A T A A T T T T G A T T G A T G C G A T G G T C A T A T T A A A T G A C T A C T T T G G C 38
QY 2775 C T C T G C A C T G G C T T C A G A T G G A C A C A C A G A T T C T A A C T 2811
D b 37 C T C T G C A C T G G C T T C A G A T G G A C A C A C A G A T T C T A A C T 1

RESULT 7

US-09-442-100-1

Sequence 1, Application US/09442100

Patent No. 6359193

GENERAL INFORMATION:

APPLICANT: Xu, Tian

APPLICANT: Tao, Wufan

APPLICANT: Wang, Weiyl

APPLICANT: Zhang, Sheng

APPLICANT: Yu, Wan

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS

TITLE OF INVENTION: GENES AND METHODS BASED THEREON

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 869-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1103..4402
US-09-442-100-1

Query Match 11.8%; Score 469; DB 4; Length 5720;

Best Local Similarity 63.0%; Pred. No. 4.1e-112;

Matches 819; Conservative 0; Mismatches 455; Indels 27; Gaps 5;

QY 2088 A C A C A A A G A T G A A G A G G A G G A A T C T C G T A T T C A A A G T A T T C T C T C A A G C A T T T 2147

D b 2984 A G G A G A G G A G G A G G A G G C A A G G A G T T C C G C A T C A G G C A G T A C T C G C G C A A G C T T C 3043

QY 2148 A A A T T C T T A T G G A G C A A C A T G T A G A A A T G T A C T A A A T C T C A T C A G C A G C G T C T A C A T 2207

D b 3044 A A C T T C T C A T G G A G C A G C A C A T A G A C A G C G A T C A A G T C G T A T C C C A G C C A C G T A T 3103

QY 2208 C G T A A A A A C A A T T A G A G A A T G A A T G C C G G T T G G A T T A T C T C A A G A T G C C C A G G A T 2267

D b 3104 C G C A A G A A T C A G C T G G A G A A G G A G A T G C A A A G T G G G A C T C C C G A T C A G A C C C A A A T C 3163

QY 2268 C A A A T G A G A A G A T G C T T T C C C A A A A A G A A T C T A A T T A C A T C C G T T T A A A A G G C G T A A 2327

D b 3164 G A G A T G A G A A A A T G C T G A A C C A A A A G G A G A G A C A C A T A C A T T C G A T T G A A G C G C C A A G 3223

QY 2328 A T G G A C A A G T C A T G T T T G T G A A G A T A A A G A C A C A T A G G A T A G G A C A T T T G G T G A A G T C 2387

D b 3224 A T G G A C A A G A C A T G T T C G T C A A C T A G A C C C A T T G G A G T G G T G C A T T T G C G A G G T A 3283

QY 2388 T G T C T A G C A A G A A A A C T A G A T A - - - C T A A G G C T T T G T A T G C A C A A A A A A A C T C T T C G A A A G 2444

D b 3284 A C G C T G T G A G C A A A A A T C G A T A C C T C G A A C C A T T T C T A T G C G A T G A A A C C C T C G C G A A A 3343

QY 2445 A A A G A T G T T C T T C G A A A T C A A T C A A G T C G C T C A T G T T A A G C G T G A G A G A T A T C C T G G C T 2504

D b 3344 G C G A C G T T C T C A A G C G G A A T C A G T G G C A C A C A G T G A A G C C G A G A G A G A T A T C C T C G G 3403

QY 2505 G A A G C T G A C A A T C A A T G G G T A G T T C G T C T A T A T T A T T C A A G A T A A G A A C A A T T A 2564

D b 3404 G A A G C C A C A A A C T G G G T G G T G A A G T T G T A C T A C A G C T T C C A G C A A G A A G T A A T C T G 3463

QY 2565 T A C T T T G T A A T G G A C T A C A T T C C T G G G G T G A T A T G A T G A C C T A T T A A T T A A T G A A T G G C 2624

3464	DB	TACTTTCTGATGGACTACATACCAGGTGGTGATCTGATCTGCTGCTCATCAACTGGCG	3523
2625	QY	ATCTTCCAGAAAGCTGGCAGGATTTACATAGCAGAACATTTACCTGTGCAGTTGAAAGT	2684
3524	DB	ATTTTCGAGGAGGAGACTGGCCAGATTTACATCGCGAGGTCACTCGCGCGTGGACAGC	3583
2685	QY	GTTCATAAAATGGGTTTTATTATAGAGATATTAAACCTGATATAATTTTGATTGATCGT	2744
3584	DB	GTTCACAAAATGGGTTTCATTTACAGAGACATCAAGCCTGACAACTACTCATCGATAGG	3643
2745	QY	GATGTCATATTAAATTTACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACAGAT	2804
3644	DB	GACGACACATAAAGCTCACCGACTTTGGCCTGTGCACGGGATTCGATGGACGACAAAC	3703
2805	QY	TCTAAGTACTATC---AGAGTGGTGACCATCCACGCCAAGATAGCATCGATTTTCAGTAAT	2861
3704	DB	TCGAAGTACTACCAGGAGACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAG	3763
2862	QY	GAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCCATTTAGACGGCGAG	2921
3764	DB	GAATCTCCGAGAAC-----GGACCGAGCCCAACCGTCTGTGGAGAGCGCA	3808
2922	QY	GCTGCACGCCACACCGACGATGCTAGCACATCTTTTGGTGGGACTCCCAATTTATATT	2981
3809	DB	CGATCGCGGATCACCAGAAGAGTCTGGCCCACTCGTGGCGCACCCCGAAGCTACATA	3868
2982	QY	GCACCTGGAAGTGTGCTACGRACAGGATACACACAGTTGTGATTGGTGGAGTGTGGT	3041
3869	DB	GCTCCCGAGGTGCTGGAGAGGAGTGGGTACACGACGTGTGCGACTACTGGAGCGTGGC	3928
3042	QY	GTTATCTTTTGAATGTGTGTGGGACAACCTCCTTTCTTGGCACCAACACCATAGAA	3101
3929	DB	GTATCTTTACGAGATGCTGTGTGGTCAAGCGCCCTTTCTGGCCAACTGTCGCTGGAA	3988
3102	QY	ACACAAATGAAGTTATCAACTGGCAACAATCTCTTCACATTCACACCAAGCTAAATC	3161
3989	DB	ACGCACAAAAGTCACTCACTGGGAGAAACGTCATATTCGCGCGCAGGCGGAGTTA	4048
3162	QY	AGTCCTCAAGCTTCTGATCTTATTAATACTTTGCCGAGACCCGAGATCCCTTTAGGC	3221
4049	DB	TCCCGGAGGCTACGGACTTGATAAGAGGCTCTGTGCGTGGCTGACACCGCGCTGGC	4108
3222	QY	AGAATGCTGCTGATGAATATAAGCTCATCCATTTTTTAAACAATTTGACTTCTCCAGT	3281
4109	DB	AGA---GCGTGGACGAGGTCAACGCCACGACTCTTCAAGGCGATCGACTTGC---G	4162
3282	QY	GACCTTGACACAGCACTCTGCTTCATACATTCCTAAATTCACACACCCCAACGATACATCA	3341
4163	DB	GACATGCCGAGACAGAAAGCGCCCTACATACCGGAATCAACGACCCCAACCGCACATCC	4222
3342	QY	AAATTTGATCTGTTGATCTCGTGAATAATTTATGGAGTGATGA	3382
4223	DB	AACTTTGATCCCGTGGATCCCGGAGCAAGCTCGCGTCCGAATGA	4263

RESULT 8
US-09-328-111-66
; Sequence 66, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

```

: TITLE OF INVENTION: PRODUCTS
:
: FILE REFERENCE: CCD-257 (US)
:
: CURRENT APPLICATION NUMBER: US/09/328,111
:
: CURRENT FILING DATE: 1999-06-08
:
: EARLIER APPLICATION NUMBER: US 60/088,801
:
: EARLIER FILING DATE: 1998-06-10
:
: NUMBER OF SEQ ID NOS: 850
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 66
:
: LENGTH: 678
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: misc_feature
:
: LOCATION: (1)...(678)
:
: OTHER INFORMATION: n = A, T, C or G
:
: S-09-328-111-66

```

Query Match	11.4%;	Score 452.8;	DB 4;	Length 678;
Best Local Similarity	90.1%;	Pred. No. 2.4e-108;		
Matches 548;	Conservative	0;	Mismatches 50;	Indels 10; Caps
QY - 2180	ACTCAAATCTCATGACGAGCGTCTACATCGTAAACAAATTTAGAGAATGAATGATGCG	2239		
DB 1	ACTCAAATCTCATGACGAGCGTCTACATCGTAAACAAATTTAGAGAATGAATGATGCG	60		
QY 2240	GGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAGATC	2299		
DB 61	GGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAGATC	120		
QY 2300	TAATTACATCCGCTTTAAAGGGCTAAATGCGACAGTCTCTTTGTGGAAGATAAAGC	2359		
DB 121	TAATTACATCCGCTTTAAAGGGCTAAATGCGACAGTCTCTTTGTGGAAGATAAAGC	180		
QY 2360	ACTAGGAATAGGAGCATTTGGTGAAGTGTCTAGCAAGAAAGTAGACTACTAAGGCTTT	2419		
DB 181	ACTAGGAATAGGAGCATTTGGTGAAGTGTCTAGCAAGAAAGTAGACTACTAAGGCTTT	240		
QY 2420	GTATGCAACAAAACHTCTTCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGT	2479		
DB 241	GTATGCAACAAAACHTCTTCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGT	300		
QY 2480	TAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACACATGAATGGTGTAGTTCGTCTATT	2539		
DB 301	TAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACACATGAATGGTGTAGTTCGTCTATT	360		
QY 2540	TTCATTCACAGATAAGACAAATTTATFACHTTTGTAATGGACTACATTCCTGGGGTGATAT	2599		
DB 361	TTCATTCACAGATAAGACAAATTTATFACHTTTGTAATGGACTACATTCCTGGGGTGATAT	418		
QY 2600	GATGAGCCTATTAAATAGAAATGGGCATC-TTTCGAGAAGCTGCGACGATTCATACAT--	2656		
DB 419	GAAGAGCCCATTAATTTANAATGGGCATCTTTCCAGAAGGGCTNGACCAATCTACCTTA	478		
QY 2657	AGCGAAGCTTACCTGTG-CAGTTTGAAGTGTTCATAAAA--TGGGTTTTATTTCATAGAGA	2713		
DB 479	GCGAGAAGCTTACCTGTGCGCCNGTTTGAAGATGGTCTTTAAAATGGGCTTTAAATCTTAGAGA	538		
QY 2714	TATTAACCTTGATATAATTTTGTATGATC--GTGATGGTCATATTTAAATGACGTGACTTT	2771		
DB 539	TTTTTAACTGGATAATTTTGTATGATGATC--GTGATGGTCATATTTAAATGACGTGACTTT	598		
QY 2772	GGCCTGTG 2779			
DB 599	GGCCTGTG 606			

RESULT 9
US-08-878-989-11
; Sequence 11, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

[illegible]

Db 845 CTTGGATTGGAGATTTTGGATCTCTTAAAGTATAGCGAGAGGAGCATTTGGTGAGGTA 904
QY 2388 TGTCTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAGAAA 2447
Db 905 CGGCTTGTTGAGAAAGAAATACGGGACATGCTGTATGCAATGAAATATACCTCGGTAAGCA 964
QY 2448 GATGTTCTTCTTCAAAATCAAGTCTCATGTTAAAGCTGAGAGAGATATCCTGGCTGAA 2507
Db 965 GATATGTTGAAAAGAGAGAGTTGGCCACATTCGTGGGAGCGTGACATCTTAGTGGAG 1024
QY 2508 GCTCAGCAATGAATGGGTAGTTCGCTATATATTATTCATCCAAAGATAAGGACATTTTATAC 2567
Db 1025 GCAGACAGTTTCTGGGTTGCAAAATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTAC 1084
QY 2568 TTTGTAATGGACTACATTCCTGGGGGTGATGATGAGCCCTATTAATAGAAATGGGGATC 2627
Db 1085 CTAATCATGAGAGTTCTGCTGGAGGGGACATGATGACCTTTGTTGATGAAAAGACACT 1144
QY 2628 TTTCAGAAAGTCTGGCAGGATTTCTACATAGCAGAACTTACCTCTGCAGTTGAAAGTGT 2687
Db 1145 CTGACAGAAAGAGAGACTCAGTTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATT 1204
QY 2688 CATAAAATGGGTTTTATTATCATAGAGATATTAACCTTGATATATTTTATGATCGTGAT 2747
Db 1205 CACCAACTTGGATTCATCCACAGAGAGATCAAAACAGACAACTTCTTTTGGACAGCAAG 1264
QY 2748 GGTCAATTAATTAAGTACTGCTTTGGGCTCTGCACTGGCTTCAGATGACACAGATTCCT 2807
Db 1265 GGCCATGTGAAACTTTCTGACTTTTGGCTTTTGCACAGACTGAAAAGAGCAGTAGGACA 1324
QY 2808 AAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGGATTTTCAGTAATGAATGG 2867
Db 1325 GAATTTTATAGAACTCAACACAGCCTCC-----CCAGTGATTTTACATTTCCAGAAC 1378
QY 2868 GGGATCCCTCAAGCTGTGATGTGGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCA 2927
Db 1379 ATGAATTCAAAAG-----GAAAGCAGAAAACCTGGA 1409
QY 2928 GCCAGCAGCAGCGATGTCTAGCACATTTCTTGGTGGGACTCCCAATATATATTTGCACT 2987
Db 1410 AAAGA-AATAGACGTCAGCTAGCCTTCCACAGTAGGACACTCCTGACTACTATTGCTCCT 1458
QY 2988 GAAAGTGTGCTACGAAACAGGATACACAGATTTGTTGATTTGGTGGAGTGTGGTGTATT 3047
Db 1469 GAGGTGTTCTACGACAGCCGGGTACAAAGCTCTGTGATTTGGTGTGCTTTGGGGTGATC 1528
QY 3048 CTTTGTGAATGTGGTGGGACAACTCTTCTTGGCACAACACCACTTAGAACACAA 3107
Db 1529 ATGTATGATGCTCATCGGCTACCCACTTCTGTTCTGAGACCCCTCAAGAGACATAT 1588
QY 3108 ATGAAGGTATCAACTGGCAACATCTCTTCACATTTCCACCAAGCTTAAACTCAGTCTCT 3167
Db 1589 AAGAAGGTATGACTGGAAGAAACTTTGACTTTTCTCCAGAGTTCCCATCTCTGAG 1648
QY 3168 GAAGCTTCTGATCTTATTAATAACTTTGGGAGAGACCCGAGATCGCTTAGGCAAGAT 3227
Db 1649 AAGGCCAAGGACTTAATTTGAGTCTCTGTGAATGGGAACATAGAAATTTGGAGCTCCT 1708
QY 3228 GTGCTGATGAATTAAGCTCATCCATTTTAAACAATTTGACTTCTCCAGTCACTG 3287
Db 1709 GAGGTGTAGGAATAAAGTAATCTCTTTTGAAGCGGTGA---CTGGGAACATATC 1765
QY 3288 AGACAGCAGCTCTGCTTATACATTTCTTAAATACACACCCCAACATACATCAAAATTTT 3347
Db 1766 AGAGAGAGACCTGCTGCAATATCTATTGAATCAAAAGCATTGATGATCACTCAAACTTC 1825
QY 3348 GAT 3350
Db 1826 GAT 1828

RESULT 13

US-09-588-256-1
; Sequence 1, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas
; APPLICANT: Flavler, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippsen, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
US-09-588-256-1

Query Match 4.3%; Score 173.2; DB 4; Length 2160;
Best Local Similarity 56.2%; Pred. No. 2.7e-35;
Matches 325; Conservative 0; Mismatches 253; Indels 0; Gaps 0;
QY 2243 TGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAACTAA 2302
Db 828 TGGTGTCTCTGAAGAGAGAAGACACAGACTTTCTTCTACCTGGGTAAAGAGAGTCGCA 887
QY 2303 TTACATCCGCTTAAAGGGCTAAATGGCAAGCTATGTTGTGAAGATAAAGACACT 2362
Db 888 GTTCTCGCTTGGTAGGACACGGCTATCCCTGGAGATTTCCACACTGTTAAAGTCAT 947
QY 2363 AGGAATAGGAGCAATTTGGTGAAGTCTCTAGCAAGAAAAGTAGATAGTACAGGCTTTGTA 2422
Db 948 AGAAAGGGTGCATTCGTGTAGGTCCTGCTGTCGAGAAAGAAAGATACCCGGTAAATATA 1007
QY 2423 TGCAACAAAACCTCTCGAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTAA 2482
Db 1008 CGCTATGAAGACATGTTTAAATCAGAAATGTACAAGAGATCAATTAGCCACGTCAA 1067
QY 2483 GGCTGAGAGAGATATCCTGGCTGAAGCTGCAATGAATGGGTAGTTCGTCTATATATTC 2542
Db 1068 GGCCGAGAGGATGTTGGCCGGAAGGACTCTCCGTGGGTCTGCTGTATATCTATTC 1127
QY 2543 ATTCAGATGAAGACAAATTTATCTTTGTAATGGACTACATTCCTGGGTGATATGAT 2602
Db 1128 TTTCAGATGCCAGTACCTATATCTTGATCATGGAATTTTGGCCGGTGGTACCTGAT 1187
QY 2603 GAGCTATTAATAGATGGCATCTTTCCAGAAAGTCTGCGACGATTTCTACATAGCAGA 2662
Db 1188 GACCATGTTATCAGTGGCAGATATTCACCGAGGACGTCACAGATTTCTACATGGCGGA 1247
QY 2663 ACTTACCTGTCCAGTGAAGTGTTCATAAATGGGTTTTTATTCATAGAGATATTAACC 2722
Db 1248 GTGTATCTCGCAATTTAGGCTATACACAAGCTGGGCTTTTATTCATAGAGATATCAAGCC 1307
QY 2723 TGATAAATTTTGTGATGCTGATGCTGTCATTAATTAATTTGACTTGGCTCTGCTGCAC 2782
Db 1308 GGATAACATTTCTGATCGACATCAGGGGTACATCAAACTTTCCGACTTCGGTCTGCTGCAC 1367
QY 2783 TGGCTTTCAGATGGACACACGATTTCTAAGTACTATCAGA 2820
Db 1368 AGGGTTCACAAAGGCATGACTCAACTACTACAAGA 1405

RESULT 14
US-08-860-150-1

Db	915	TTTATAGGCACGTGCCCCACAGTCCGATGGACTCCAAGCGACGTGCCGAGTCGTGGAACCGA	974
Qy	2914	AGCGGAGAGCTGCACCGACGACGAGCGATGCTTAGCACATCTTTTGGTTGGGACTCCCA	2973
Db	975	AATCGACGCCCTCGCTACAGCAC-----GTGGGAACGCCGG	1014
Qy	2974	ATTATATPGCACTGAAGTGTGCTACGAAACAGGATACACACAGTGTGTGATTTGGTGGGA	3033
Db	1015	ACTATATPGCAACCGAAGTATTCTGCAGACTGGCTACGGACCCGCGCTGCGACTGGTGGT	1074
Qy	3034	GTGTTGGTGTATTCTTTTGTGAATGTTGGTGGGACACACTCTCTTCTTGGCACAACAC	3093
Db	1075	CCCTGGGAGTCATCATGTACGAAATGCTGATGGGCTATCTCTCCATTTCTGTCGGACAATC	1134
Qy	3094	CATTAGAAACACAAATCAAGGTTTATCAACTGCGCAACATCTCTTCACATTCACCCACAAG	3153
Db	1135	CCAGGACACTACCCAGAGGTGATGAACTGGCGGACAGCGTGATATTTCCTCCCAAG	1194
Qy	3154	CTAAACTCAGTCTGAAGCTTCTGATCTTATTATTAAACTTTGCGGAGGACCCGAAGATC	3213
Db	1195	ATCCCATATCGGAGGAGGCCAAGGAGACGATCATCAACTTCTGTGCGAGGCCGATCGCC	1254
Qy	3214	GCATTAGCCAAAGTGTGCTGATGATGAATATAAAGCTCATCTATTTTAAACAATTTGACT	3273
Db	1255	GCTGGGTTCCAGCGTCTCTGGAGGATCTGAAAGTCGTGCCGTTCTTCGGGGGAGTTGACT	1314
RESULT 15			
US-09-338-132-1			
; Sequence 1, Application US/09338132			
; Patent No. 6040164			
; GENERAL INFORMATION:			
; APPLICANT: Hemmings, Brian A.			
; APPLICANT: Millward, Thomas A.			
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases			
; FILE REFERENCE: 4-20265/A/PCT			
; CURRENT APPLICATION NUMBER: US/09/338,132			
; CURRENT FILING DATE: 1999-06-22			
; EARLIER APPLICATION NUMBER: 08/860,150			
; EARLIER FILING DATE: 1997-06-19			
; EARLIER APPLICATION NUMBER: PCT/EP95/05052			
; EARLIER FILING DATE: 1995-12-20			
; EARLIER APPLICATION NUMBER: 94810746.1			
; EARLIER FILING DATE: 1994-12-22			
; NUMBER OF SEQ ID NOS: 18			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 1			
; LENGTH: 2101			
; TYPE: DNA			
; ORGANISM: Drosophila melanogaster			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (132)...(1499)			
US-09-338-132-1			
Query Match 4.3%; Score 172.4; DB 3; Length 2101;			
Best Local Similarity 49.9%; Pred. No. 4.3e-35;			
Matches 539; Conservative 0; Mismatches 511; Indels 30; Gaps			
Qy	2197	AGCGTCTACATCGTAAAAAACAAATTAGAAGATGAATGATGCGGTTGGATTATCTCAAG	2256
Db	262	AGCGAAGACGCCCTCGCAAGCTGGAGGCTCACTGAAGGACGAGAGCTTGTGCGAGG	321
Qy	2257	ATGCCCAAGATCAAAATGAGAAAGATGCTTTGGCAAAAAAGATCTAATTACATCGGCTTTA	2316
Db	322	CGCAGGCCACAGGAGAAGCGCTCTGCACATGTCGCCAGAAGGACGAGTATCTCGCGTGA	381
Qy	2317	AAAGGCTTAAATGGCAAGTCTATTGTTGTGAAGATAAAGACACTAGGATAGGACAT	2376
Db	382	AGCATTTGCGCCTCGGTGTGGAGGACTTTTGGAGGCCCTCAAAAGTCATCGGACGCGCGCT	441
Qy	2377	TTGGTGAAGTCTCTGTACGCAAGAAAGTAGATACTAAAGCTTTTGTATGCAACAAAACTC	2436

```

Db 442 TCCTGGAAGTCGCTTTGGTCAGAAAAAGGACACTGGACATGTGTGCGCCATGAAGGTGC 501
Qy 2437 TTCGAAAGAAAGATGTTCTTCTTCGAATCAAGTCGCTCATGTTAAGCTGAGAG---AG 2493
Db 502 TGCACAAAGCGGACATGCTGGAAGAGGAGCAGAGTGGCAGACGTACGCGCGGAGGGTCTGC 561
Qy 2494 ATATCCTGGCTGAGCTGACAAATGAATGGGTAGTTCGTCTATATTTATTTCCCAAGATA 2553
Db 562 ATGCTCTGGTCGAGCGCGATCATCATGTTGGTGGTGAAGATGTACTACAGTTTCCAGGATC 621
Qy 2554 AGGACAAATTTATCTTTGTAATGGACTACATTCCTGGGGGTGATATGATGAGCCCTATTAA 2613
Db 622 CCGTCAATTTATATTGATAATGAGTCTTGCCTGGTGGTGATGATGACGCTTTTAA 681
Qy 2614 TTAGAAATGGGATCTTTCCAGAAAGTCTGGCAGCATCTACATAGCAGAACTTACCTGTG 2673
Db 682 TGAAGAAGGACACGCTATCCGAGGAGGACACAGTTCTATATCAGTGAGACGGCATGG 741
Qy 2674 CAGTTGAAAGTGTTCATAAAATGGTATTATTCATAGATATTTAAACCTGATAATATTT 2733
Db 742 CGATCGATTCTATTCAAACTCGTTTCATACAGGGATATCAAGCCGATTAACCTGC 801
Qy 2734 TGATGTGATCGTGTGATATTAATTAAGTACTTTGGCTCTGCACTGGCTTTCAGAT 2793
Db 802 TGTGGACGCGGAGGCGATCTGAAGCTCTCCGACTTCGGACTGTGCACTGGCTTAAAGA 861
Qy 2794 GGACACAGATTTCTAAGTACTATCAGAGTGTGACCATCCAGCGCAAGATAGCATGGATT 2853
Db 862 AGTCGATCGAACAGACTTTTATCGGGACTTGTGCGAGC-----GAAACCATCCGAT 914
Qy 2854 TCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCATTAG 2913
Db 915 TTTATAGGCACGTGGCCAGTCCGATGGACTCCAAAGCGACGTGCGAGTCTGGAAGCGA 974
Qy 2914 AGCGGAGAGCTGCAGCGCAGCAGCAGCATGTCTAGCACATTTCTTGGTTGGGACTCCCA 2973
Db 975 AATCGACGCGCCCTCGCCTACAGCACC-----GTGGGAACGCCGG 1014
Qy 2974 ATTATATTGCACTCAAGCTTGTGTACGACAGGATACACAGTTGTGATTGGTGCA 3033
Db 1015 ACTATATTGCAACCGAAGTATTTCTGCAGACTGGCTACGACCGCCCTGCGACTGGTGGT 1074
Qy 3034 GTGTTGGTGTATTCTTTTGAATGTTGGTGGGACAACTCTCTTTTGGCACAACAC 3093
Db 1075 CCCTGGGAGTCATCATGTACGAAATGCTGATGGGCTATCCTCCATTCGCTCGGACAATC 1134
Qy 3094 CATTAAGAACACAAATGAAGGTTATCAACTGGCAACATCTCTTCACATTTCCACCACAAG 3153
Db 1135 CCCAGGACACCTACCGCAAGGTGATGAATGGCGGAGACGCTGATATTTCCCCCAAGAG 1194
Qy 3154 CTAAACTCAGTCTCAAGCTTCTGATCTTATTATTAACCTTTCGAGGAGCCCGAAGATC 3213
Db 1195 ATCCCATATCGGAGGAGGCGCAAGGAGAGCATCATCACTTCTGTCGAGGCGCATGCC 1254
Qy 3214 GCTTAGGCAAGAAATGGTCTGATGAATAAAGCTCATCCATTTTAAAAACAATGTACT 3273
Db 1255 GCTGGGTTCCAGCGCTCTGGAGGATCTGAAGTCTGCCGTTCTTCCGGGGAGTTGACT 1314

```

Search completed: January 16, 2003, 21:55:17
Job time : 162.058 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: January 16, 2003, 09:56:42 ; Search time 110.556 Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: 1.accttgggtgctgggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 39868 seqs, 22934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	655.4	16.5	2043	9	US-09-836-392-2
2	609.8	15.3	638	10	US-09-879-536-26
3	452.8	11.4	678	10	US-09-879-536-66
4	346.2	8.7	676	9	US-09-764-868-214
5	261.4	6.6	1689	9	US-09-938-842A-1861
6	254.6	6.4	3583	9	US-09-974-298-152
7	216.8	5.4	1452	9	US-09-938-842A-2402
8	186	4.7	734	9	US-09-764-868-196
9	171.8	4.3	1818	10	US-09-771-161A-89
10	122.8	3.1	2706	12	US-10-071-751-20
11	120.6	3.0	1416	9	US-09-938-842A-2503
12	120.2	3.0	5313	10	US-09-801-368-297
13	119.2	3.0	1398	9	US-09-938-842A-633
14	117	2.9	568	9	US-09-796-692-8122
15	117	2.9	6409	10	US-09-864-864-293
16	116.2	2.9	2637	10	US-09-799-875-3
17	116.2	2.9	3003	10	US-09-799-875-1
18	112.6	2.8	2365	9	US-09-981-353-6
19	112.4	2.8	277	10	US-09-294-093B-2043

20	112.4	2.8	1244	10	US-09-771-161A-38	Sequence 38, Appl
21	112.4	2.8	1393	10	US-09-771-161A-37	Sequence 37, Appl
22	111.6	2.8	1735	9	US-09-764-868-58	Sequence 58, Appl
23	111	2.8	2311	10	US-09-810-808-6	Sequence 6, Appl
24	111	2.8	2370	9	US-10-000-039-1	Sequence 1, Appl
25	111	2.8	2370	10	US-09-969-347-214	Sequence 214, App
26	111	2.8	2370	10	US-09-880-107-3855	Sequence 3855, Ap
27	107.4	2.7	587	9	US-09-764-868-217	Sequence 217, App
28	106.4	2.7	1194	10	US-09-801-368-399	Sequence 399, App
29	105.2	2.6	1479	10	US-09-771-161A-46	Sequence 46, Appl
30	101.4	2.5	1515	10	US-09-804-471A-1	Sequence 1, Appl
31	101.4	2.5	5877	12	US-10-028-946-3	Sequence 3, Appl
32	101.4	2.5	6165	12	US-10-028-946-1	Sequence 1, Appl
33	100.4	2.5	2261	9	US-09-954-531-141	Sequence 141, App
34	100.4	2.5	2261	9	US-09-954-531-570	Sequence 570, App
35	99	2.5	1197	10	US-09-801-368-403	Sequence 403, App
36	98.2	2.5	2346	10	US-09-954-456-1137	Sequence 1137, Ap
37	96.6	2.4	6159	9	US-10-017-216-3	Sequence 3, Appl
38	96.6	2.4	6574	9	US-10-017-216-1	Sequence 1, Appl
39	95	2.4	2512	10	US-09-784-249-1	Sequence 1, Appl
40	93.4	2.3	1143	10	US-09-801-368-401	Sequence 401, App
41	93.4	2.3	2760	9	US-10-098-841-195	Sequence 195, App
42	92.6	2.3	2549	10	US-09-880-107-3691	Sequence 3691, Ap
43	92.2	2.3	1470	9	US-09-938-842A-1978	Sequence 1978, Ap
44	91.4	2.3	362	9	US-09-796-692-7789	Sequence 7789, Ap
45	91.4	2.3	3061	10	US-09-880-107-2146	Sequence 2146, Ap

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match	16.5%	Score 655.4;	DB 9;	Length 2043;
Best Local Similarity	69.2%	Pred. NO. 1.6e-148;		
Matches '909;	Conservative 0;	Mismatches 401;	Indels 3;	Gaps 1;
OY	2241	GTGATATCTCAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAAATCT	2300	
Db	218	GCTGAGCTGTGTGAAGCTGAGCAGGAGATGCGGAAGATCCTCTACGAAAGAGTCT	277	
OY	2301	AAATACATCCGCTCTTAAAGGGCTAAATGGAACAAGTCTATCTTTGTGAAGATAAGACA	2360	
Db	278	AATTACACAGCTTAAAGAGGCCCAAGATGGACAGTCTATCTTTGTCAAGATCAAAACC	337	
OY	2361	CTAGGAATAGGAGCATTTGGTGAAGTCTCTCTAGCAAGAAAAGTAGACTAAGGCTTTG	2420	
Db	338	CTGGGGATCGGTGCTCTTTGGAGAAGTGTGCTTGTGTAAGGTGGACACTCACGCCCTG	397	

QY 2421 TATGCAACAAAACTCTTCCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTT 2480
 DB 398 TAGCCCATGAAGACCCCTAAGGAAAAAGAGTGTCTGAAACCGGAATCAGGTGGCCACGTC 457
 QY 2481 AAGCTGAGAGATATCTCTGGCTGAGCTGACAATGAATGGTGTCTGCTCTATATAT 2540
 DB 458 AAGCCGAGAGGACATCTCTGGCGGAGGACAAATGAGTGGGTGTCTCAAACTCTACTAC 517
 QY 2541 TCATTCCAGATAGGCAATTTATATCTTTGTAATGAGCTACATCTCTGGGGTGATG 2600
 DB 518 TCCTTCCAGACAAGACACCCCTGACTTTGTGTGATGACTACATCCCTGGTGGGACATG 577
 QY 2601 ATGAGCCTATTAAATGAGATGGGATCTTTCCAGAAAGTCTGCAGCATCTACATAGCA 2660
 DB 578 ATGAGCCTGTGTATCCGGATGGAGGTCTTCCCTGAGCACCTGGCCGGTCTTACATCGCA 637
 QY 2661 GAACCTTACCTGTGAGTGAAGTGTTCATAAAATGGTTTTTATCATAGAGATATATAA 2720
 DB 638 GAGCTGACTTTGGCCATTGAGAGTGTCCACAAGATGGCTTCATCCACCGAGACATCAAG 697
 QY 2721 CCTGATAATATTTGATGCTGATGGTGGTCAATATTAATGACTGTGCTTGGCCTCTGC 2780
 DB 698 CTTGATACATTTTGTATGATCTGCTGATGCTGATTAATTAACCTACAGATTTTGGCCTCTGC 757
 QY 2781 ACTGGCTTCAGATGACACACGATTTCTAAGTACTATCAGAGTGTGACCATCCACGGCAA 2840
 DB 758 ACTGGCTTCAGTGGACTCAAAATTCACAAATTTACCAAGAAAGGACCACTGTCAGACAG 817
 QY 2841 GATAGCTGATTTTCAGTAATGAATGGGGGATCCCTCAAGCTGTGATGTTGGAGACAGA 2900
 DB 818 GACAGCTGAGACCCAGGACCTCTGGATGATGTGCTAATCTGCTGCTGCTGGGACAGG 877
 QY 2901 CTGAAGCCATTAGAGCGGAGCTGACGCGCAGCACAGCGATGCTAGACATTTCTTG 2960
 DB 878 CTGAAGCCCTTAGAGCAGAGGCGCGGAAGCAGCAGCAGAGGTGCTTGGCAGATTTCTG 937
 QY 2961 GTTGGACTCCCAATATATTTGCACTTCAAGTGTGCTTACCAAGAGGATACACACAGTTG 3020
 DB 938 GTGGGACTCCAAACTATCATCGACCCGAGGTGCTCTCCGCAAGGATGACTCAACTC 997
 QY 3021 TGTGATTTGGTGGAGTGTGTTATTTCTTTTGAATGTTGGTGGGACACCTCTCTTC 3080
 DB 998 TGTGACTGTTGGAGTGTGATTTCTCTTCGAGATGCTGGTGGGACAGCCGCCCTTT 1057
 QY 3081 TTGGCACAACACCATTTAGAACACAAATGAAGTTATCAACTGCAACATCTCTTAC 3140
 DB 1058 TTGGCACCCTACTCCACAGAACCCAGCTGAAGGTGATCACTGGGAGAACACGCTCCAC 1117
 QY 3141 ATTCCACCACAAGCTAAACTCAGTCTCCTGAAGCTTCTGATCTTTATTTAAACTTTGCCGA 3200
 DB 1118 ATTCCAGCCAGGTGAGCTGAGCTTCTGAGCCAGGACCTCATCACCAGCTGTGCTGC 1177
 QY 3201 GGACCCGAGATGCTTGGAGCAAGATGCTGATGATGAATAAAGCTCATCCATTTTTT 3260
 DB 1178 TCCGAGACACCCGCTGGGCGGAAATGGGCGGATGACCTGAGGAGCCACCCCTCTTTC 1237
 QY 3261 AAAACAATTTGACTTCTCCAGTCACTGAGACAGCAGTCTGCTTCTATACATCTCTAAATC 3320
 DB 1238 AGCGCATTTGACTTCTCCAGTGACATCCGGAAGCAGCAGCCCTACGTTCCCAACCATC 1297
 QY 3321 ACACACCCCAACAGATACATCAAAATTTTGTATCTGTTGATGCTGATTAATTTATGAGTAT 3380
 DB 1298 AGCCACCCCATGACACCTTCGAATTTCCGACCCGCTAGATGAAGAAGGCCCTTGGAAAGAT 1357
 QY 3381 GATAAGCAGAGAAATGTAATGACACTCTCAATGGATGTAATAAATGGAAGAT 3440
 DB 1358 GGCAGC- --GAAGGTAGCACCAGGCGCTGGGACACACTCACCTCGCCCAATCAAGAT 1414
 QY 3441 CCTGAACATGCTTCTATGAATTTTACCTTCCGAAGGTTTTTTTGTGACAAATGGCTACCCA 3500
 DB 1415 CTTGACACGCATTTTACGAATTCACCTTCCGAGGTTCTTTTGTGACAAATGGCTACCC 1474

QY 3501 TATAATTATCCGAGCCCTATTGAATATATATATATATATATATATATATATATATAT 3553
 DB 1475 TTTTCGATCCCAAGACCTTCAGGAGCAGAAAGCTTCACAGGCTGAGAGCTCAGA 1527
 RESULT 2
 US-09-879-536-26/c
 ; Sequence 26, Application US/09879536
 ; Patent No. US20020144298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Maria E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; TITLE OF INVENTION: PRODUCTS
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/879,536
 ; CURRENT FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/088,801
 ; PRIOR FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FASTSEQ for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-879-536-26
 Query Match 15.3%; Score 609.8; DB 10; Length 638;
 Best Local Similarity 98.4%; Pred. No. 9.2e-138;
 Matches 627; Conservative 0; Mismatches 7; Indels 3; Gaps 1;
 QY 2178 GTACTCAAACTCATCAGCAGCGCTACATCGTAAACAAATAGAGATGAATGATG 2237
 DB 637 GTACTCAAACTCATCAGCAGCGCTACATCGTAAACAAATAGAGATGAATGATG 578
 QY 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAGAA 2297
 DB 577 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAGAA 518
 QY 2298 TCTAATTACATCCGCTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAG 2357
 DB 517 TCTAATTACATCCGCTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAG 458
 QY 2358 ACATAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATTAAGGCT 2417
 DB 457 ACATAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATTAAGGCT 398
 QY 2418 TTGTATGCAACAAAACCTCTCGAAGAAAGATGCTTCTTCCGAAATCAAGTCGCTCAT 2477
 DB 397 TTGTATGCAACAAAACCTCTCGAAGAAAGATGCTTCTTCCGAAATCAAGTCGCTCAT 338
 QY 2478 GTTAAGGCTGAGAGATATCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGCTATAT 2537
 DB 337 GTTAAGGCTGAGAGATATCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGCTATAT 278
 QY 2538 TATTCAATCCAGATAAGGACAATTTATCTTTGTATGGACTACATTCCTGGGGGTGAT 2597
 DB 277 TATTCAATCCAGATAAGGACAATTTATCTTTGTATGGACTACATTCCTGGGGGTGAT 218
 QY 2598 ATGATGAGCCCTATTAAATTAGAATGGCATCTTCCAGAAAGCTTGGCAGGATTTCTACATA 2657
 DB 217 ATGATGAGCCCTATTAAATTAGAATGGCATCTTCCAGAAAGCTTGGCAGGATTTCTACATA 158

Qy 2658 GCAGAACTACCTGTCAGTGAAGTG---TTTCAAAAATGGGTTTTTATTATAGAGAT 2714
Db 157 GCAGAACTACCTGTCAGTGAAGTGTTCCTTAAATGGGTTTTTATTATAGAGAT 98
Qy 2715 ATTAACCTGATAATATTTTTCATTCATGTCATGTCATATTAATTAATGACTTTGGC 2774
Db 97 ATTAACCTGATAATATTTTTCATTCATGTCATGTCATATTAATTAATGACTTTGGC 38
Qy 2775 CTCTGCACTGGCTTCAGATGGACACAGCATTTCTAAGT 2811
Db 37 CTCTGCACTGGCTTCAGATGGACACAGCATTTCTAAGT 1

RESULT 3

US-09-879-536-66
; Sequence 66, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(678)
; OTHER INFORMATION: n = A,T,C or G
US-09-879-536-66

Query Match 11.4%; Score 452.8; DB 10; Length 678;
Best Local Similarity 90.1%; Pred. No. 8.2e-100;
Matches 548; Conservative 0; Mismatches 50; Indels 10; Gaps 6;
Qy 2180 ACTCAAACTCATCAGCAGCGCTACATCGTAAACAAATAGAGAAATGAATGATCG 2239
Db 1 ACTCAAACTCATCAGCAGCGCTACATCGTAAACAAATAGAGAAATGAATGATCG 60
Qy 2240 GGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAAGAATC 2299
Db 61 GGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAAGAATC 120
Qy 2300 TAATACATCCGTCTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAAGAC 2359
Db 121 TAATACATCCGTCTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAAGAC 180
Qy 2360 ACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGACTAAGGCTTT 2419
Db 181 ACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGACTAAGGCTTT 240
Qy 2420 GTATGCACAAAACACTCTTCAAGAAAGATGTTCTTCTTGAATCAAGTCGCTCATGT 2479
Db 241 GTATGCACAAAACACTCTTCAAGAAAGATGTTCTTCTTGAATCAAGTCGCTCATGT 300

Qy 2480 TAAGCTGAGAGAGATATCCCTGGCTGAAGCTACAAATGAATGGGTAGTTCGCTATATTA 2539
Db 301 TAAGCTGAGAGAGATATCCCTGGCTGAAGCTACAAATGAATGGGTAGTTCGCTATATTA 360
Qy 2540 TTCATTCCAAGATAAGGACAATTTTATCTTTGTAATGAGCTACATTCCTGGGGTGATAT 2599
Db 361 TTCATTCCAAGATAAGGAGCCATTTATCCTT--GTAATGGCTACATTCCTTNGGGTGATAT 418
Qy 2600 GATGAGCCTATTAATAGAAATGGGCATC-TTTCAGAAAGCTGGCAGATTTCTACAT-- 2656
Db 419 GAAGAGCCCATTAATANAATGGGCATCTTTTCCAGAAAGGCTNGCACAATCTACCTTA 478
Qy 2657 AGCAGAACTTACCTGTG-CAGTTGAAAGCTTTCATAAAA--TGGGTTTTTATTTCATAGAGA 2713
Db 479 GCCAGAACTTACCTGNGCCNGTTGAAAGTGGTCTTAAATGGGGTTTAAATCTTAGAGA 538
Qy 2714 TATTAACTGATAATATTTGATGATGTC--GTGATGTGCATATTAATTAATGACTGACTTT 2771
Db 539 TTTTAACTGATAATATTTGANTGGACCGNAGAGGCGCTTATTAAAAATGGCTTGCTTT 598
Qy 2772 GGCCTCTG 2779
Db 599 GGCCTTNG 606

RESULT 4

US-09-764-868-214
; Sequence 214, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (628)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-214

Query Match 8.7%; Score 346.2; DB 9; Length 676;
Best Local Similarity 71.2%; Pred. No. 4.7e-74;
Matches 481; Conservative 2; Mismatches 191; Indels 2; Gaps 2;
Qy 2145 TTTAAATTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTA 2204
Db 1 TTTAAGTTCTTCATGGAGCAGCAGCTGGAGAATGTATCAAAAACCTACCAGCAGAAAGTT 60
Qy 2205 CATCGTAAAAACAATTAGAAGTAAGATGATGCGGGTTGGATTATCTCAAGATGCCAG 2264
Db 61 AACCGAGGCTGACGTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120
Qy 2265 GATCAATAGCAAGACATGCTTTGCCAAAAGAATCTAATATACATCCGCTTTAAAAGGCT 2324
Db 121 GAGCAGATGCGGAGATCTCTACCAGAAAGAGTCTAATACACAGAGTTAAAGAGGCC 180
Qy 2325 AAAATGGACAAGTCTATGTTGTGAAGATAAGACACTAGGAATAGGAGCATTTGGTGAA 2384
Db 181 AAGATGGACAAGTCTATGTTGTCAAGATCAAAACCTGGGATGGTGGCTTTTGGAGAA 240
Qy 2385 GTCTGCTCTAGCAAGAAAAGTAGACTAAGGCTTTGTATGCAACAAAACACTCTTCGAAG 2444
Db 241 GTGTCCTTGTGTAAGTGGACACTACGCGCTGTACGCCATGAAGACCTTAAGGAAA 300
Qy 2445 AAGATGTTCTTCTTCGAAATCAAGTCGCTCATGTGTTAAGGCTGAGAGAGATATCCTGCT 2504

Db 301 AAGATGTCTGACCGAATCAGTGGCCAGTCAAGCCGAGAGGACATCTGGCC 360
 QY 2505 GAAGCTGACAAATGAATGGGTGCTCTATATATTTATTTCAATTCAGAGTAAAGGACAAATTA 2564
 Db 361 GAGGACAGACAAATCAGTGGGT-GGCAAACTCTACTACTCTCTTCCAAAGACAAAGACGCTG 419
 QY 2565 TACTTTGTAATGACATACATTCCTGGGGTGATATGATGAGCCTATTAATAGAAATGGG 2624
 Db 420 TACTTTGTAATGACATACATTCCTGGGGTGATATGATGAGCCTATTAATAGAAATGGG 479
 QY 2625 ATCTTCCAGAAAGTCTGGCAGCAATCTTACATAGCAGAACTTACCTGTGCGAGTTGAAAGT 2684
 Db 480 GTCTTCCCTGAGCAGCTGGCGGCTCTACATCGCA-RACTGACTTTGGCCTATGAGAGT 538
 QY 2685 GTTCATAAAATGGGTTTATTCATAGAGATTAATAACCTGATATAATTTGATGATCGT 2744
 Db 539 GTCCACAAGATGGGCTTATCCACCGAGACATCAAGCCTGATGAATTTGATAGATCTG 598
 QY 2745 GATGGTCATATAAATGACTGACTTTGGGCTCTGCACTGGCTTCAGATGGACACAGAT 2804
 Db 599 GATGGTCATATAAATGACTGACTTTGGGCTCTGCACTGGCTTCAGATGGGCTTCAGTGGACTCACAT 658
 QY 2805 TCTAAGTACTATCAGA 2820
 Db 659 TTCCAATATTACCAGA 674

RESULT 5

US-09-938-842A-1861
 : Sequence 1861, Application US/09938842A
 : Patent No. US20020160378A1
 : GENERAL INFORMATION:
 : APPLICANT: Harper, Jeff
 : APPLICANT: Kreps, Joel
 : APPLICANT: Wang, Xun
 : APPLICANT: Zhu, Tong
 : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 : TITLE OF INVENTION: SAME, AND METHODS OF USE
 : FILE REFERENCE: SCRIP1300-3
 : CURRENT APPLICATION NUMBER: US/09/938, 842A
 : CURRENT FILING DATE: 2001-08-24
 : PRIOR APPLICATION NUMBER: US 60/227,866
 : PRIOR FILING DATE: 2000-08-24
 : PRIOR APPLICATION NUMBER: US 60/264,647
 : PRIOR FILING DATE: 2001-01-16
 : PRIOR APPLICATION NUMBER: US 60/300,111
 : PRIOR FILING DATE: 2001-06-22
 : NUMBER OF SEQ ID NOS: 5379
 : SEQ ID NO 1861
 : LENGTH: 1689
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 : US-09-938-842A:1861

Query Match 6.6%; Score 261.4; DB 9; Length 1689;
 Best Local Similarity 52.7%; Pred. No. 2.4e-53;
 Matches 642; Conservative 0; Mismatches 566; Indels 11; Gaps 3;
 QY 2142 GCATTTAAATTTTATGGGACACATGTAGAAAATGTACTCAATCTCATCAGGCGT 2201
 Db 157 GCAGCGAAGCAATATATCGAGAATCATATAAAATTCAGAGAAGAGTCTCCAGGAAGA 216
 QY 2202 CTACATCTTAAAAACAATTAGAAGTGAATGATGGGTTGGATTATCTCAAGATGCC 2261
 Db 217 AAAGAGCTCGGAGCACTTCGACAAACACCTAGCTGATGCTGATGTTGTAAGAC 276
 QY 2262 CAGGATCAAAATGAGAAAGATGCTTTGGCAAAAGAACTTAATTCACGCTCTTAAAGG 2321
 Db 277 AAGATGGATATTAAGAACTTTGAGAAAAAGGAAATGGAGTATATGCGTCTACAAGA 336
 QY 2322 GCTAAATGGCAAGTCTATCTTCTGAGATAAGACACTAGGAATAGGAGCAATTTGGT 2381

Db 337 CAGAAAATGGGGTTGATGACTTTGAACCTGCTTAAGCATCATTTGGCGGGGCTTCTTCGGTT 396
 QY 2382 GAAGTCTGCTAGCAAGAAAAGTAGATACCTAAGGCTTTGTATATGCAACAAAACTCTTCGA 2441
 Db 397 GAGGTGAGAAATTTGTAAGAAAAATCTACTGGAAGCTATATGCAATGAAAAGTTAAAG 456
 QY 2442 AAGAAAGATGTTCTTCTTCGAAATCAAGTCGCTCATGTTAAAGCTGAGAGATATPCCTG 2501
 Db 457 AAATCCGAGATGCTTCGACGAGGAGGTGAACATGTTAAAGCTGAAAAGAAATGTCTT 516
 QY 2502 GCTGAAGCTGACAAATGAATGGGTAGTTCCTCTATATTTATTTATTTCCCAAGATAAGACAT 2561
 Db 517 GCAGAACTGATAGCCCTTTCATGTCAGCTTTGTTACTCTTTCCCAAGATGATGACAT 576
 QY 2562 TTATATCTTTGTAATGACTACATTTCTGGGGTGATATGATGAGCCTATTAATTAAGAATG 2621
 Db 577 TTGTATCTTTATGGAATACCTCCCTGGAGTGATATGATGACATGCTGATGCGAAG 636
 QY 2622 GGATCTTTCCAGAAAGTCTGGACGATTTCTACATAGAGAACTTACCTGTGCAAGTGA 2681
 Db 637 GATACCTTACGGAAGATGAGACTCGTTTTATTTGTCGACAGCAATTCGCTGATTTAG 696
 QY 2682 AGTGTTCATAAAATGGTTTTATTTCATAGAGATATTAAACCTGATATAATTTTATTGAT 2741
 Db 697 TCTATCCATAAGCAATAATTAGTCCACAGGATATAAGCCTGATATAATTTATTGATTA 756
 QY 2742 CGTGTGTCATATAATAATTTGACTGACTTTGGCCTCTGCACTGGCTTCAGATGGACAC 2801
 Db 757 CGAAACGGCCATATCAAGCTTTTCAGATTTTGGATTGAGCAAGTCTCTGGAAGCAAAA 816
 QY 2802 GATTTAAGTACTATCAGAGTGTGACCATCCAGCGCAAGATAGCATGATTTCACTAAT 2861
 Db 817 TTTCCAGATTTTAA-----GGCGAGGCTTTGTGACAGGAGTACAAAGCCTCAGCAGA 869
 QY 2862 GAATGGGGGATCCCTCAAGCTCTCGATGTGGAGAC-AGACTGAAGCCATTAGAGCGGAG 2920
 Db 870 ACATGATAGACTCTCAAGCCTCTCTGCACTAGAGAACTCAGGAGAACAGCTTTT 929
 QY 2921 AGCTGACCGCCAGCACAGCGATGTCTAGCAATCTTTGGTTGGGACTCCCAATATAT 2980
 Db 930 ACATTGGCAACAAACAGAAAGGACCTCGCTTTTCTACAGTAGGAATCCCGGATTA 989
 QY 2981 TGCACTGAAGTGTGCTAGCAAGACAGATACACAGTTGCTGTGATTGGTGGAGTGTGG 3040
 Db 990 TGCCCTGAGGTGCTGCTGAAGAAAGGATGGAATGGAGTGTGATTGGTGGTCTCTGG 1049
 QY 3041 TGTATTCTTTTGAATGTTGGTGGGCAACACCTCTCTTTCTGGCACAACACCAATTAGA 3100
 Db 1050 AGCAATCATGTTGAGATGCTTGTAGGTTTCCGCCATCTATTCCGAAGAACCTTTGG 1109
 QY 3101 AACACAAATGAAGTTATCAACTGGCAACATCTCTTCACATTCACCACCAAGCAACT 3160
 Db 1110 AACATGTAGAAAGATTGTAACCTGGAACCTGCTTGAAGTTCCCTGATGAAGCTAGCT 1169
 QY 3161 CAGTCTGAAGCTCTGATCTTATTAATTAACCTTTGCGGAGGACCCGAGATCGCTTAGG 3220
 Db 1170 CTCATGCGAGTAAAGATCTTATCCGAAAGCTGCTGCAATGTCGAACAGAGGCTTGG 1229
 QY 3221 CAAGAATGGTCTGATGAAATAAAAGCTCATCCATTTTAAACAAATGACITCTCCAG 3280
 Db 1230 GACCAAGAGATTCAAGAAATCAAGCAGACACCTCTGTTGTTAGGGGAGTCCGAATGG 1289
 QY 3281 TGACCTGAGACAGAGTCTGCTTCATACATCTCTTAATCAGACACCCACACATACATC 3340
 Db 1290 ---GCTATATGAGTCAAAATGCTCCATATATACCACAAGTGAAGCATGAGCTTGATAC 1346
 QY 3341 AAATTTGATCTCTGAT 3359
 Db 1347 AAATTTGAAAGTTTGTAT 1365

Sequence 152, Application US/09974298
Patent No. US20020156263A1
GENERAL INFORMATION:
APPLICANT: Chen, Hwei-Mei
TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
FILE REFERENCE: PA-0037 P
CURRENT APPLICATION NUMBER: US/09/974,298
CURRENT FILING DATE: 2001-10-04
PRIOR APPLICATION NUMBER: 60/238,331
PRIOR FILING DATE: 2000-05-10
NUMBER OF SEQ ID NOS: 194
SOFTWARE: PERL Program
SEQ ID NO 152
LENGTH: 3583
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. US20020156263A1 347975.11
US-09-763-298-152

Query Match 6.4%; Score 254.6; DB 9; Length 3583;
Best Local Similarity 53.6%; Pred. No. 1.6e-51;
Matches 645; Conservative 0; Mismatches 519; Indels 39; Gaps 4;

QY 2148 AATTCTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTACAT 2207
DB 353 AAGTGACACTGGAGAAATTTTATGCAACCTTATGCTCAACATGAAGAACGAGAAATG 412

QY 2208 CGTAAAAACAATAGAGAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGAT 2267
DB 413 AGACAAAGAAAGTTAGAAAAGGTGATGGAAGAGGCGCTAAAGAGATGAGGAGAAACGA 472

QY 2268 CAATGAGAAAGATGCTTTGCCAAAAGAAATCTAATATACATCGCTTTAAAGGGCTAA 2327
DB 473 CTCGGAGATGACACATGCTCGGAAGAGAAACAGAGTTCTTCGTTTGAAGAGAACAGA 532

QY 2328 ATGGCAACTCTATGTTGTGAAGATAAAGACACTAGGAATAGAGCAATTTGGTGAATGC 2387
DB 533 CTGTGATTGGAAGATTTTGTAGTCCGTAAAGTAATAGGAGGAGCAATTTGGTGAAGTA 592

QY 2388 TGCTAGCAAGAAAGTAGATCTAGGCTTTGTATGCAACAAAACACTCTTCGAAGAAA 2447
DB 593 CGGCTTTGTCAGAGAGAGATCGGACATGTGTATGCAATGAAATATACTCGTAAAGCA 652

QY 2448 GATGTTCTTCTCGAAATCAAGTCGCTCATGTAAGGCTGAGAGATATCTGCTGAA 2507
DB 653 GATATGCTTGAAGAGAGAGGTTGGCCACATTCGGCGGAGCGTGACATTTCTAGTGGAG 712

QY 2508 GCTGCAATGAATGGGTAGTCTGTCTATATTATTCATTCGAAGATAAGGACAAATTTATAC 2567
DB 713 GCAGACAGTTTCTGGGTTGTGAATAATGTTCTATAGTTTTCAGATAAGCTAAACCTCTAC 772

QY 2568 TTTGTAATGACTACATCTCTGGGGTGATATGATGAGGCTATTAATTAAGATGGGCATC 2627
DB 773 CTAATCATGGAGTTCTGCTGGAGGGGACATGATGACCTTTGTATGAAAAAGACACT 832

QY 2628 TTTCCAGAAAGTCTGCGAGATTCTACATAGCAGAACTTACCTGTGCACTTTGAAAGTGT 2687
DB 833 CTGACAGAGAGAGACTCAGTTTATATAGCAGAAACAGTATATAGCCATGACTCTATT 892

QY 2688 CATAAAATGGGTTTATTCATAGAGATATTAACCTTGATAATATTATTGATTGCTGTAT 2747
DB 893 CACCAACTTGGATTATCCACAGAGACATCAAAACAGACAACTTCTTTTGACAGCAAG 952

QY 2748 GGTATATTAAATGACTGACTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2807
DB 953 GGCCATGTGAAACTTTCTGACTTTGCTTTTGCACAGGACTGAAAAAGACATAGGACA 1012

QY 2808 AAGTACTATCAGAGTGGTCAACATCCACGGCAAGATAGATGATGATTTTCAATGATG 2867
DB 1013 GAATTTATAGGAATCTGAACCAACAGCCCTCC-----CCAGTGATTTTCTTCCAGAAC 1066

QY 2868 GGGGATCCCTCAAGCTGTGCGATGTGGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCA 2927
DB 1067 ATGAATTCACAAAG-----GAAGCAGAACTGGA 1097

QY 2928 CGCCAGCACCGAGATGTCTAGCACATCTTTGGTTGGAGCTCCCAATATATATGCACT 2987
DB 1098 AAAGA-AATAGAGCGTACAGCTAGCTTCTCCACAGTAGGCACCTCTGACTACATTTGCTCT 1156

QY 2988 GAAGTCTTCTAGACAGAGATACACAGATTGTGTGATTGGTGGAGTGTGGTGTATT 3047
DB 1157 GAGGTGTTTCATGAGACCGGTTACACAAAGCTCTGTGATTGGTGGTGGTGGTGGTATC 1216

QY 3048 CTTTTTGAATGTGTGGGACAACTCTCTTTTGGCAACAAACACCACTTAGAACAACAA 3107
DB 1217 ATGTATGAGATGCTCATCGCTTACCACCTTCTGTCTGAGACCCCTCAAGAGACATAT 1276

QY 3108 ATGAAGGTTATCACTGGCAACATCTCTTACATCTTCCACCAAGCTTAACTCAGTCTCT 3167
DB 1277 AAGAAGGTGATGAACCTGGAAAGAACTTTTGACTTTTCTCCAGAAAGTTCCCATCTCTGAG 1336

QY 3168 GAAGCTTCTGATCTTATTATTAACCTTTCCCGAGGACCGGAGAGATCGCTTAGCAAGAT 3227
DB 1337 AAGCCAAAGATCTAATTTTGGAGTCTGCTGTGAATGGAAACATAGAATTTGAGCTCTCT 1396

QY 3228 GGTGCTGATGAATAAAAGCTCATCCATTTTAAACAAATTTCTCTCCAGTGACCTG 3287
DB 1397 GGAGTTGAGGAATAAAAGTAACCTTTTGTGAGGCGTTGA---CTGGGAACATATC 1453

QY 3288 AGACAGACTCTGCTTCTATACATCTCTTAAATCAACACACCCCAAGATACATCAATTTT 3347
DB 1454 AGAGAGAGACCTGCTGCTCAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTC 1513

QY 3348 GAT 3350
DB 1514 GAT 1516

RESULT 7
US-09-938-842A-2402
Sequence 2402, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2402
LENGTH: 1452
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2402

Query Match 5.4%; Score 216.8; DB 9; Length 1452;
Best Local Similarity 50.4%; Pred. No. 1.3e-42;
Matches 583; Conservative 0; Mismatches 567; Indels 6; Gaps 2;

QY 2220 TTAGAGATGAATGATCGGGTTCGATTTATCTCAAGTCCCGCAGCATCAATGACAAAG 2279
DB 157 TTAGAGAGAAAGTTCGGCTCTTCTGGAGTGCAGAGAGGAGGAGCAATCAACATGATATA 216

QY 2280 ATGCTTTGCCAAAAGAAATCTAATTAATACATCCGCTTTAAAGGGCTTAAATGACAAAGTCT 2339

Db 217 GACCTGGAGCAAGAGACAGAGCTTTATGAGCTTAAAGGAACAAGATTAGTGTGAT 276
 Qy 2340 ATGTTTGAAGATAAAGACACTAGGAATAGGACCAATTTGGTGAAGTCTGTCTAGCAAGA 2399
 Db 277 GACTTTGAGCTTTTGAAGTATCATTTGAGAGAGGCTTTTGGTGAAGTCTGTCTAGCGG 336
 Qy 2400 AAAGTAGATCAAGGCTTTGATGCAACAACTCTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 337 GAGAAAGCTGGAATATTTATGCAATGAAGATTAAGAAATCTGAATGCTCATG 396
 Qy 2460 CGAATCAAGTCGCTATGTTAAGGCTGAGAGATATCTTGGCTGAGCTGACAATGAA 2519
 Db 397 AGAGCAGGTTGAGCATGTGAGAGCAGAGAGAACTGCTGGCTGAGGTTGAAGCCAT 456
 Qy 2520 TGGGTAGTCTCTATATTTATTCATTCAGATTAAGACAAATTTATACCTTTGTAATGAC 2579
 Db 457 TATATTGGAAGCTTTACTATTTCAATTTCCAGATCCGAGTATCTGTATCTGATTATGAA 516
 Qy 2580 TACATCTCTGGGCTGATATGAGCTATTAATTAAGATGGCCTTTTCCAGAAAGT 2639
 Db 517 TATCTCCCGGTGGTATGATGACCTTGTCTCATGAGGAGGACACATACGGGAAGAT 576
 Qy 2640 CTGGCAGCATTTACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATPAATAATGGT 2699
 Db 577 GTTCCAGATTTTATTTGCTCAAGTCTCTGGCCATTTGAATCCATACAGATACAAC 636
 Qy 2700 TTTATTCATAGATATTAACCTGATTAATTTGATGATGCTGATGCTGATTAATAA 2759
 Db 637 TATATTTCATAGGATATCAAACTTGAATTAATTTTGGACAAAGATGGCCATGAA 696
 Qy 2760 TTGACTGACTTTGGCTCTGCACTGGCTTCAGATGGACACAGATTTCAAGTACTATCAG 2819
 Db 697 CTCTGGACTTTGGCTCTGTAAGCTCTGATTTGTAAGAAATTTACCTTCAATTCAGGAG 756
 Qy 2820 AGTGGTG--ACCATCCAGGCAAGATAGCATGGAATTTCAAGTAATTAATGCGGAGTCCCT 2877
 Db 757 AATAGGGCCAGGATGATGAACCTATGTCAGAACCTATGATGATGATGATGATGCTTCT 816
 Qy 2878 CAAGCTGTCGATGGAGACAGACTGAGCCATTTAGCGGAGAGCTGCA---CGCCAG 2933
 Db 817 GACACTGATACAGAGAAAGCTGGCGCAGTCCCGAGAACAACTTCAGACTGGCAGATG 876
 Qy 2934 CACCAGCGATGTAGCAGATCTTTGGTGGGACTCCCAATTAATTTGCACTGAAAGTG 2993
 Db 877 AATCCAGAAACTAGCATTTTCAACTGTGGAAACACCGGACTATATGCTCTCTGAAAGTT 936
 Qy 2994 TTGCTACGACAGATACACAGTTGTTGATTTGGTGGGAGTGTGGTGTATTTCTTTT 3053
 Db 937 TTGCTGAAGAAGGATATGGCATGGAATGTGATTGGTGGTCTATTAGTGTCAATTTATGTAC 996
 Qy 3054 GAAATGTTGGTGGGCAACCTCTTTCTTGGCACAACACCACTTAGAACACCAATGAAG 3113
 Db 997 GAAATGCTGTTGGTATCTCTCTTTTATGCTGATGACCTATATCACTTGCAGAAAG 1056
 Qy 3114 GTTATCAACTGGCAACATCTCTTCAATTTCCACCAAGCTAACTCAGTCTGAAAGCT 3173
 Db 1057 ATCGTCCATTTGAGAAACCAATTTGAAATTTCTGAGGATGCGAAGTTTCTATCTGAGGCA 1116
 Qy 3174 TCTGATCTTATTAACCTTTGCGAGGACCCGAGATGCTTAGGCAAGAAATGGTCT 3233
 Db 1117 AAGATCTTATCTCAGGTTGCTGTGCAACGTTGACCATAGGCTTGGTACTGGAGGAGGA 1176
 Qy 3234 GATGAATTAAGCTCATCTCTTTTAAACAAATTTGACTTCTCCAGTGACCTGAGACAG 3293
 Db 1177 GCCAGCAATCAAGGATCATCTTGGTTCAAGATGTTGTTGGGAAAGCTCTATGAA 1236
 Qy 3294 CAGTCTGCTCATACATTTCTTAAATACACACCCCAAGATACATCAATTTTGTATCCT 3353
 Db 1237 ATGAGGCTGCTACAAACCAAGAGTGAACGACGAGCTAGATACACAAATTTTATGAAG 1296
 Qy 3354 GTTGATCTGATAAT 3369

Db 1297 TTTGATGAAGTGAATT 1312
 RESULT 8
 US-09-764-868-196
 : Sequence 196, Application US/09764868
 : Patent No. US20020168711A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 : FILE REFERENCE: PIZ32
 : CURRENT APPLICATION NUMBER: US/09/764,868
 : CURRENT FILING DATE: 2001-01-17
 : Prior application data removed - refer to PALM or file wrapper
 : NUMBER OF SEQ ID NOS: 1510
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 196
 : LENGTH: 734
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-764-868-196

Query Match 4.7%; Score 186; DB 9; Length 734;
 Best Local Similarity 55.6%; Pred. No. 2.5e-35;
 Matches 419; Conservative 2; Mismatches 297; Indels 36; Gaps 2;
 Qy 2400 AAAGTAGATCAAGGCTTTGATGCAACAACTCTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 12 AAGAAATACAGCCCATATCTATGCAATGAAGATATTGAGAAAGTCTGATATGCTTGA 71
 Qy 2460 CGAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTGCTGAGAGTGAAGTGA 2519
 Db 72 AAGAGCAGGTTGGCCCATATCCGAGCAGAAAGAGATATTTGGTARAAGCAGATGGTCC 131
 Qy 2520 TGGTAGTCTCTATATTTATTTCCAGATAAGGACATTTATCTTTGTAATGAC 2579
 Db 132 TGGTGTGAGATGTTTACAGTTTTCAGGATAAGAGGAATCTTTATCTANTCATGAA 191
 Qy 2580 TACATCTCTGGGCTGATGATGAGCTATTAATTAAGATGGCCTCTTCCAGAAAGT 2639
 Db 192 TTTCTCCCTGGAGTGACATGACATGCTAATGAAGAAAGACWCCTTTCAGACAGAG 251
 Qy 2640 CTGCGAGATCTACATAGCAGAACTTACCTGCTGAGTGAAGTGTTCATAAATGGT 2699
 Db 252 GAAACACAGTTTACATTTTCAGAGACTGTTCTGCAATAGATGCGATCCACAGTGGGT 311
 Qy 2700 TTTATTCATAGATATTAACCTGATAATTTTGTGATGCTGATGCTGATGCTATATAA 2759
 Db 312 TTCACTCATCGGATATTAGCCAGACACCTTTTATGATGCCAAGGTCATGATAA 371
 Qy 2760 TTGACTGACTTTGGCTCTGCACTGGCTTCAGATGGACACAGATTTAAGTACTATCAG 2819
 Db 372 TTATCTGATTTGGTTTATGTACGGGATTAAGAAAGCTCACAGGACTGAATTTTATAGA 431
 Qy 2820 AGTGGTGACCATCCAGCAGATAGCATGATTTCACTATGATGATGGGGGATCCCTCA 2879
 Db 432 AATCTCACACACACCCCAAGT-----G 456
 Qy 2880 AGCTGTGCTGAGACAGACTGAAGCCATTAGAGGGAGAGCTGCGACGCCACCCAG 2939
 Db 457 ACTTCTCATTTCAACATGACTCAAGAGAAAGCAGAACTTGGAG--AAGAACAGG 515
 Qy 2940 CGATGCTAGCAGATCTTTGGTGGGACTCCCAATTTATTTGACCTGAGTGTGCTA 2999
 Db 516 AGACAACCTGGCATATCCACAGTTGGGACACAGATTACATTTGCCAAGAAATTTATG 575
 Qy 3000 CGACAGATACACACAGTCTGTGATTTGGTGGAGTGTGTGTTTATTTTGAATG 3059
 Db 576 CAGACTGTTTACACAAATTTGTGACTGTGGTCTTTGGGAGTCATATGATGAAATG 635
 Qy 3060 TTGTTGGGACACCTCTTCTTGGCACAACACCACTTGAAGACACAAATGAAGTTATC 3119
 Db 636 CTAATAGGATATCCACCTTTCTGCTCTGAAACACCTCAAGAGACATACAGAAAGTATG 695

Fri Jan 17 11:16:55 2003

Db 215 GCATTTTGGGAAGACGTCATATAATGCTCATGCAAAATCAGAAATGATGTACAA 274
 Qy 2532 CTATATTATTCATCCAGACAAATTTATACCTTTGTAATGGACTACATTCCTGGG 2591
 Db 275 TTACATTTTGGTTTTCAGAGATCAAAATATCTTATATGTCATGGATATATGCCGGG 334
 Qy 2592 GGTGATATGATGACCTTAAATAGAAATGGCATCTTTCCAGAAAGTCTGGCAGATTC 2651
 Db 335 GGTGACTTGGTGAGTCCTTATGTCGATATGAAA---TTCCAGAAAAATGGCAATGTC 391
 Qy 2652 TACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATAAAATGGTCTTATTCATAGA 2711
 Db 392 TATACAATGGAAGTGTGTCACACTTATACAAATCTCCATGGGATTTGTACATCGT 451
 Qy 2712 GATATTAACCTGATAATATTTTGTATGATCGTATGCTCATATTAATTAATGACTGACTTT 2771
 Db 452 GATGTTAAACCTGATAATATGCTTCTAGACAAATATGCTCATTTTAAAGTTAGCTGACTTT 511
 Qy 2772 GGCCTCTGCA 2781
 Db 512 GGAACCTGTA 521

RESULT 11
 US-09-938-842A-2503
 ; Sequence 2503, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPT1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2503
 ; LENGTH: 1416
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2503

Query Match: 3.0%; Score 120.6; DB 9; Length 1416;
 Best Local Similarity 55.0%; Pred. No. 2.3e-19;
 Matches 237; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
 Qy 2343 TTTGTGAAGATAAAGACACTAGGAATAGGACATTTGGTGAAGTCTGTAGCAAGAAA 2402
 Db 418 TTTGAGTATTGAAGTTTGGGAGAGGTGCAATTTGGAAAAGTTACCAGGTGAGGAAA 477
 Qy 2403 GTAGATCTAAGGCTTTGTATGCAACAAACCTTCGAAGAAGATGTTCTTCGA 2462
 Db 478 AAAGACACGCTGTGATATACGCGATGAAGGTCTAGAGAAAGATAAAATGTTGAGAAG 537
 Qy 2463 AATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTGCGTGAAGCTGACAATGAATG 2522
 Db 538 AATCATGCTGAATCATCAAGCCGAGCGGATATTTAACCAAAATCGATCATCTCTTC 597
 Qy 2523 GTAGTCTGCTATATTTATTCATTCAGATGAAGCAATTTATACCTTTGTAATGACTAC 2582
 Db 598 ATTGTGCAACTTAAATACCTTTTTCAGACCAAAATACAGATGATGTTGTTCTTGACTTT 657
 Qy 2583 ATTCCTGGGGTGTATGATGATGAGGCTTATTAATAGAAATGGGCATCTTTCCAGAAAGCTG 2642

Db 658 ATAAACGGAGGTGTCATCTTTCTTCCAGCTCTATCACCAGGGCTTTTCAGGAGGACTTG 717
 Qy 2643 GCACGATTTCTACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGTGT 2702
 Db 718 GTCGCTGTGTACACTCAGAAAATCGTCTCTGAGTTTCCCATCTCTCAGTGAAGGCATA 777
 Qy 2703 ATTCTATAGATATTAACCTGATATAATTTTGTATGATCGTGGTGCATATTAATTTG 2762
 Db 778 ATCCATAGAGATCTTAACCTGAAACATACATCATGACGTAGATGGCCATCTGATGTTA 837
 Qy 2763 ACTGACTTTGG 2773
 Db 838 ACTGATTTGG 848

RESULT 12
 US-09-801-368-297
 ; Sequence 297, Application US/09801368
 ; Patent No. US20020128250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Busby, Robert
 ; APPLICANT: Cali, Brian
 ; APPLICANT: Hecht, Peter
 ; APPLICANT: Holtzman, Doug
 ; APPLICANT: Madden, Kevin
 ; APPLICANT: Maxon, Mary
 ; APPLICANT: Milne, Todd
 ; APPLICANT: No. US20020128250A1man, Thea
 ; APPLICANT: Royer, John
 ; APPLICANT: Salama, Sofie
 ; APPLICANT: Sherman, Amir
 ; APPLICANT: Silva, Jeff
 ; APPLICANT: Summers, Eric
 ; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 ; FILE REFERENCE: 109272.147
 ; CURRENT APPLICATION NUMBER: US/09/801,368
 ; CURRENT FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: US 09/487,558
 ; PRIOR FILING DATE: 2000-01-19
 ; PRIOR APPLICATION NUMBER: US 60/160,587
 ; PRIOR FILING DATE: 1999-10-20
 ; NUMBER OF SEQ ID NOS: 440
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 297
 ; LENGTH: 5313
 ; TYPE: DNA
 ; ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-297

Query Match: 3.0%; Score 120.2; DB 10; Length 5313;
 Best Local Similarity 56.6%; Pred. No. 5.8e-19;
 Matches 243; Conservative 0; Mismatches 183; Indels 3; Gaps 1;
 Qy 2353 TAAAGACACTAGGAATAGAGAGATTTGGTGAAGTCTGTAGCAAGAAAAGTAGATACTA 2412
 Db 2390 TGAACCAATATCAGCAAGAGTCTTATGTTAGTGTATCTAGCAGGAAAAAACTCAG 2449
 Qy 2413 AGCTTTGTATCCACAAACACTCTCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCG 2472
 Db 2450 GAGATATTTTGTCTATAAAGGTTCTAAGGAATCAGATATGATGCCCAAAATCAAGTAA 2509
 Qy 2473 CTCATCTTAAGGCTGAGAGAGATATC---CTGCGTGAAGCTGACAATGAATGGTAGTTC 2529
 Db 2510 CAATGTCAAATCCGAGAGAGCAATCATGTTGTTCAAAAGTATAGCCCTATGTTGGA 2569
 Qy 2530 GTCATATATTATTCATCCAGATAGGACAATTTACTTTCTTAATGGAATCAATTCCTG 2589
 Db 2570 GACTATTTGCTAGTTTCCAAAATAAAGATAACCTTTCTTAGTGATGATATTTACCAG 2629
 Qy 2590 GGGTGTATGATGAGCCCTTAAATTAATAGATGGCATCTTTCCAGAAAGCTGCGCAGAT 2649
 Db 2630 GTGAGATTTGGCCACTTTTAATCAAGATGATGGGTATCTGCCGATCAATCAATGGCAAGC 2689

QY 2653 ACATAGCAGAACTTACCTCTGCAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGAG 2712
DB 1035 AACTCTGCTGAAGTTGTTCTTCTGCTGCAATACACTCCATGGGTTTAAATACACAGAG 1094
QY 2713 ATATTAACTGATAATATTTTGAATGATCGTGCATGCTCATATTAATTAATGACTGCTTTG 2772
DB 1095 ATGTGAAGCTGACAAACATGCTCTTGGATAAAACATGGACATCTAAAAATTAGCAGATTTG 1154
QY 2773 GCCTCTGCA 2781
DB 1155 GCACGTGTA 1163

Search completed: January 16, 2003, 22:02:57
Job time : 183.556 secs

DB 206 TTTTATGCTTTCAAGATGATAGGTATCTGTACATGGTAATGGAGTACATGCTCGGTG 265
QY 2593 GTGATATGATGAGCTATTAATTAGAAATGGCATCTTTCCAGAAAGTCTGGCAGATTC 2652
DB 266 GAGACCTTGTAAACC---TTATGAGTAATTATGATGCTGAAATAATGGCCAAATTT 322
QY 2653 ACATAGCAGAACTTACCTCTGCAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGAG 2712
DB 323 AACTCTGCTGAAGTTGTTCTTCTGCTGCAATACACTCCATGGGTTTAAATACACAGAG 382
QY 2713 ATATTAACTGATAATATTTTGAATGATCGTGCATGCTCATATTAATTAATGACTGCTTTG 2772
DB 383 ATGTGAAGCTTGACAAACATGCTCTTGGATAAAACATGGACATCTAAAAATTAGCAGATTTG 442
QY 2773 GCCTCTGCA 2781
DB 443 GCACGTGTA 451

RESULT 15
US-09-864-864-293
; Sequence 293; Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darrick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864.864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 293
; LENGTH: 6409
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-864-293

Query Match 2.9%; Score 117; DB 10; Length 6409;
Best Local Similarity 56.2%; Pred. No. 3.8e-18;
Matches 241; Conservative 0; Mismatches 185; Indels 3; Gaps 1;
QY 2353 TAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACTA 2412
DB 738 TAAAGTTTATGGAAGAGGTGCTTTTGGTGAAGTGCAGTTGCTGCACAGGCATCGC 797
QY 2413 AGCCTTTGTATGCACAAACAACTCTTCGAAGAAAGATGTTCTTCTCGAAATCAAGTCG 2472
DB 798 AGAAGGTTTATGCTATGAAGCTCTTAGTAAGTTTGAATGATAAAAGATCAGATTCTG 857
QY 2473 CTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAAATGAATGGTGTGCTC 2532
DB 858 CCTTTTTTTGGGAAGAAAGAGATATTATGCGCTTTGGCCAAATAGCCCTGGTGGTTGAGC 917
QY 2533 TATATTATTCATCCAAAGATAGGACAATTTATCTTTGTAATGGACTACATTCCTGGGG 2592
DB 918 TTTTATGCTTCAAGATGATAGGTATCTGTACATGTAATGGAGTACATGCTGCTGGTG 977
QY 2593 GTCATATGATGACCTATTAAATTTAGAAATGGCATCTTTCCAGAAAGTCTGGCAGCATCT 2652
DB 978 GAGACCTTGTAAACC---TTATGAGTAATTATGCTGCTGAAATAATGGCCAAATTTT 1034

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 3443.86 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-1

Perfect score: 3984

Sequence: 1 accttgggttgctggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0.

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: em_estba.*

2: em_esthum.*

3: em_estin.*

4: em_estnu.*

5: em_estov.*

6: em_estpl.*

7: em_estro.*

8: em_hic.*

9: gb_estl.*

10: gb_est2.*

11: gb_hic.*

12: gb_est3.*

13: gb_est4.*

14: gb_est5.*

15: em_estfun.*

16: em_estom.*

17: gb_gss.*

18: em_gss_hum.*

19: em_gss_inv.*

20: em_gss_pln.*

21: em_gss_vrt.*

22: em_gss_fun.*

23: em_gss_mam.*

24: em_gss_mus.*

25: em_gss_other.*

26: em_gss_pro.*

27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	930.8	23.4	1807	11 AF119846	AF119846 Homo sapi
C 2	804.8	20.2	1216	11 BC015665	BC015665 Homo sapi
C 3	697.4	17.5	721	9 AI114863	AI114863 HAL1474 Hu
C 4	650.8	16.3	694	10 AV708169	AV708169 AV708169
C 5	590.6	14.8	657	14 BM971435	BM971435 UI-CF-D01
C 6	560	14.1	1203	12 BG331628	BG331628 602432628

C 7	559.8	14.1	637	9 AA910802	AA910802 ol48d06.s
C 8	555.2	13.9	578	10 BE093083	BE093083 RC5-BT074
C 9	548.2	13.8	597	14 BQ309028	BQ309028 WR0-BT450
C 10	543.2	13.6	550	10 AW502306	AW502306 UI-HF-BR0
C 11	533.4	13.4	855	10 BE568367	BE568367 601341939
C 12	530.6	13.3	673	10 BE557112	BE557112 BB557112
C 13	518.4	13.0	648	13 BI155504	BI155504 602903902
C 14	511	12.8	658	12 BG088986	BG088986 ut60g11.y
C 15	497.6	12.5	665	17 AZ340286	AZ340286 1M0072J17
C 16	495.2	12.4	555	12 BE723541	BE723541 193387 MA
C 17	491.6	12.3	667	10 BB174219	BB174219 BB174219
C 18	474.2	11.9	858	12 BE973820	BE973820 601680625
C 19	466.8	11.7	488	10 BE093084	BE093084 RC5-BT074
C 20	449.8	11.3	453	10 BE093054	BE093054 RC5-BT074
C 21	439.6	11.0	457	13 BI034906	BI034906 0V2-NN200
C 22	436.4	11.0	650	13 BI933046	BI933046 603344592
C 23	432.6	10.9	474	13 BG958580	BG958580 PM2-CT026
C 24	424	10.6	1079	13 BM460659	BM460659 AGENCOURT
C 25	414.4	10.4	600	12 BG081357	BG081357 H3064C02-
C 26	414.4	10.4	602	14 BQ201853	BQ201853 UI-R-D01-
C 27	414.4	10.4	621	12 BG081236	BG081236 H3062G12-
C 28	403.2	10.1	571	10 BE395243	BE395243 601311110
C 29	401.4	10.1	496	9 AI594372	AI594372 vol15f09.y
C 30	387.4	9.7	930	14 BQ989848	BQ989848 AGENCOURT
C 31	384.6	9.7	405	10 BE093082	BE093082 RC5-BT074
C 32	381.8	9.6	403	10 BE093081	BE093081 RC5-BT074
C 33	375.4	9.4	737	14 BM950668	BM950668 UI-M-EH04
C 34	367	9.2	376	10 BE088344	BE088344 CMI-BT068
C 35	364.4	9.1	390	14 BQ312644	BQ312644 MK0-BN011
C 36	361.8	9.1	391	14 BQ291582	BQ291582 PM3-AN002
C 37	359	9.0	559	10 AW642458	AW642458 cml19a02.w
C 38	349.8	8.8	473	9 AI430221	AI430221 me57h09.y
C 39	346.2	8.7	435	10 BB751204	BB751204 BB751204
C 40	341.4	8.6	458	9 AI596344	AI596344 me57h09.x
C 41	333.4	8.4	689	13 BI331257	BI331257 602982912
C 42	332	8.3	407	9 AA592233	AA592233 vol15f09.f
C 43	326.6	8.2	863	10 BE620135	BE620135 601483053
C 44	324.2	8.1	570	12 BF396283	BF396283 UI-R-BS2-
C 45	324.2	8.1	875	13 BI664597	BI664597 603290249

ALIGNMENTS

RESULT 1
AF119846/C
LOCUS AF119846 1807 bp mRNA linear HTC 08-MAY-2001
DEFINITION Homo sapiens PRO1474 mRNA, complete cds.
ACCESSION AF119846
VERSION AF119846.1 GI:7770128
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1807)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1807)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999) Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
FEATURES
source
1..1807
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB5626"

```

/ tissue_type="liver"
/ dev_stage="fetus"
complement(841..1530)
/ note="predicted protein of HQ1474"
/ codon_start=1
/ product="PRO1474"
/ protein_id="AAF69600.1"
/ db_xref="GI:770129"
/ translation="MADKLSRVEFKASYTFOEVHVFIOGLPEGLMAGLQVPEVQNL
IDWFFFLMCLVDHALAGSEVCRVQEGLODGLKELKEGWVLEVLQDQSVGFQI
TGFGLRLYLALQGVQVLRVPVDTLEQLAHVGSLLKHKHDFECRLPQGLKAKWR
RRRWPGVKGRRSLRPLGRALARGSPVRLPALPHOGYLGLLGLRTLAASASALRPVA
AVQELCSPLQS"
BASE COUNT      452 a   440 c   451 g   464 t
ORIGIN
Query Match      23.4%; Score 930.8; DB 11; Length 1807;
Best Local Similarity 99.8%; Pred. No. 4.1e-228;
Matches 932; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CTTTGGTTGCTGGGACGACTCTGGCCGCTCAGCGTCCGCTCAGCGCCGCTGCGCCG 62
DB 934 CTTGGGTTGCTGGGACGACTCTGGCCGCTCAGCGTCCGCTCAGCGCCGCTGCGCCG 875
QY 63 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTAAAGAAATTTAAC 122
DB 874 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTAAAGAAATTTAAC 815
QY 123 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTGTCTCAGAGAAAGCTCTGGATC 182
DB 814 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTGTCTCAGAGAAAGCTCTGGATC 755
QY 183 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAAGAGGAGT 242
DB 754 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAAGAGGAGT 695
QY 243 GAAAGCCAGAGGATATAGACAAATAGGCGCTTAAGACCTTTCCTCCAGTAACTATATCT 302
DB 694 GAAAGCCAGAGGATATAGACAAATAGGCGCTTAAGACCTTTCCTCCAGTAACTATATCT 635
QY 303 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 362
DB 634 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 575
QY 363 TCTGATGCTGCTAAGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 422
DB 574 TCTGATGCTGCTAAGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 515
QY 423 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTGCAGAAATTCGAAACTCT 482
DB 514 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTGCAGAAATTCGAAACTCT 455
QY 483 CTGCTTCCATTTGCAATGAACAAATTTCTTCGGAGTACTTCAGAGTTAATCCACAA 542
DB 454 CTGCTTCCATTTGCAATGAACAAATTTCTTCGGAGTACTTCAGAGTTAATCCACAA 395
QY 543 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATATGTTATACAAGCTCTTTCAG 602
DB 394 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATATGTTATACAAGCTCTTTCAG 335
QY 603 AAAACTTAACAAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 662
DB 334 AAAACTTAACAAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 275
QY 663 GATCCTCGAGGAGAGAGATGCTCGAGCAGCTGCAGACCTATTATATCCAGCATGAAA 722
DB 274 GATCCTCGAGGAGAGAGATGCTCGAGCAGCTGCAGACCTATTATATCCAGCATGAAA 215
QY 723 CCAGGAATGTGAGCAATCAGTTAAACCCAAACAGAGCTGGAAGGTTCTTAAAGAAATCC 782
DB 214 CCAGGAATGTGAGCAATCAGTTAAACCCAAACAGAGCTGGAAGGTTCTTAAAGAAATCC 155
QY 783 TTAGTCTCTCAGAGGATGCGCCGCCACTAGGAGAAAGTGTGCGCTTATCTATCTGAGAGT 842

```

```

DB 154 TTAGTCTCTCAGAGGCTAGGCGCCACTAGGAGAAAGTGTGGCCCTATCATCTGAGAGT 95
QY 843 CCCAACTCAGACAGATGTAGGAGACCTTTGCTGATCTGATGATATATCAGCATTTGTT 902
DB 94 CCCAACTCAGACAGATGTAGGAGACCTTTGCTGATCTGATGATATATCAGCATTTGTT 35
QY 903 CAAGCTCACCCTAGCAACGAGAGAGTGAACC 936
DB 34 CAAGCTCACCCTAGCAACGAGAGAGTGAACC 1
RESULT 2
LOCUS BC015665 1216 bp mRNA linear HTC 09-OCT-2001
DEFINITION Homo sapiens, Similar to LATS (large tumor suppressor, Drosophila)
homolog 1, clone IMAGE:4550430, mRNA.
ACCESSION BC015665
VERSION BC015665.1 GI:15990533
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1216)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-remail.nih.gov
Tissue Procurement: DCTD/DPF/Gazdar
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center (NISC),
Gaithersburg, Maryland.
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@hgrl.nih.gov
Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 33 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10862687
This clone has the following problem: incomplete processing.
Location/Qualifiers
1. 1216
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550430"
/tissue_type="Lung, large cell carcinoma"
/clone_lib="NIH_MGC_18"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
BASE COUNT. 415 a 270 c 251 g 280 t
ORIGIN
Query Match 20.2%; Score 804.8; DB 11; Length 1216;
Best Local Similarity 90.4%; Pred. No. 9.7e-196;
Matches 911; Conservative 0; Mismatches 2; Indels 95; Gaps 1;
QY 3 CTTTGGTTGCTGGGACGACTCTGGCCGCTCAGCGTCCGCTCAGCGCCGCTGCGCCG 62

```

```

/ tissue_type="liver"
/ dev_stage="fetus"
complement(841..1530)
/ note="predicted protein of HQ1474"
/ codon_start=1
/ product="PRO1474"
/ protein_id="AAF69600.1"
/ db_xref="GI:770129"
/ translation="MADKLSRVEFKASYTFOEVHVFIOGLPEGLMAGLQVPEVQNL
IDWFFFLMCLVDHALAGSEVCRVQEGLODGLKELKEGWVLEVLQDQSVGFQI
TGFGLRLYLALQGVQVLRVPVDTLEQLAHVGSLLKHKHDFECRLPQGLKAKWR
RRRWPGVKGRRSLRPLGRALARGSPVRLPALPHOGYLGLLGLRTLAASASALRPVA
AVQELCSPLQS"
BASE COUNT      452 a   440 c   451 g   464 t
ORIGIN
Query Match      23.4%; Score 930.8; DB 11; Length 1807;
Best Local Similarity 99.8%; Pred. No. 4.1e-228;
Matches 932; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CTTTGGTTGCTGGGACGACTCTGGCCGCTCAGCGTCCGCTCAGCGCCGCTGCGCCG 62
DB 934 CTTGGGTTGCTGGGACGACTCTGGCCGCTCAGCGTCCGCTCAGCGCCGCTGCGCCG 875
QY 63 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTAAAGAAATTTAAC 122
DB 874 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTAAAGAAATTTAAC 815
QY 123 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTGTCTCAGAGAAAGCTCTGGATC 182
DB 814 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTGTCTCAGAGAAAGCTCTGGATC 755
QY 183 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAAGAGGAGT 242
DB 754 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAAGAGGAGT 695
QY 243 GAAAGCCAGAGGATATAGACAAATAGGCGCTTAAGACCTTTCCTCCAGTAACTATATCT 302
DB 694 GAAAGCCAGAGGATATAGACAAATAGGCGCTTAAGACCTTTCCTCCAGTAACTATATCT 635
QY 303 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 362
DB 634 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 575
QY 363 TCTGATGCTGCTAAGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 422
DB 574 TCTGATGCTGCTAAGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 515
QY 423 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTGCAGAAATTCGAAACTCT 482
DB 514 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTGCAGAAATTCGAAACTCT 455
QY 483 CTGCTTCCATTTGCAATGAACAAATTTCTTCGGAGTACTTCAGAGTTAATCCACAA 542
DB 454 CTGCTTCCATTTGCAATGAACAAATTTCTTCGGAGTACTTCAGAGTTAATCCACAA 395
QY 543 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATATGTTATACAAGCTCTTTCAG 602
DB 394 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATATGTTATACAAGCTCTTTCAG 335
QY 603 AAAACTTAACAAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 662
DB 334 AAAACTTAACAAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 275
QY 663 GATCCTCGAGGAGAGAGATGCTCGAGCAGCTGCAGACCTATTATATCCAGCATGAAA 722
DB 274 GATCCTCGAGGAGAGAGATGCTCGAGCAGCTGCAGACCTATTATATCCAGCATGAAA 215
QY 723 CCAGGAATGTGAGCAATCAGTTAAACCCAAACAGAGCTGGAAGGTTCTTAAAGAAATCC 782
DB 214 CCAGGAATGTGAGCAATCAGTTAAACCCAAACAGAGCTGGAAGGTTCTTAAAGAAATCC 155
QY 783 TTAGTCTCTCAGAGGATGCGCCGCCACTAGGAGAAAGTGTGCGCTTATCTATCTGAGAGT 842

```

```

Db 98 CTTGGGTTGCTGGGAGGACTCTGGCCCTCAGGCTCCGCTCAGGCCGCTGGCCGC 157
QY 63 TCTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAATTTATATATGTAAGAATTTTAA 122
Db 158 TCTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAATTTATATATGTAAGAATTTTAA 217
QY 123 ACTCTGGGAGCTCTCTTGAAGGATCATTTTTCACATTTTCTCAGAGAAGCTCTGGATC 182
Db 218 ACTCTGGGAGCTCTCTTGAAGGATCATTTTTCACATTTTCTCAGAGAAGCTCTGGATC 277
QY 183 TATCAATAAAGAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGAGT 242
Db 278 TATCAATAAAGAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGAGT 337
QY 243 GAAAAGCCAGAGGATATAGACAAATGAGCCTTAAGCCTTTTCTCCAGTAATCTACT 302
Db 338 GAAAAGCCAGAGGATATAGACAAATGAGCCTTAAGCCTTTTCTCCAGTAATCTACT 397
QY 303 GTCAAGTCCAGGCAATGTTTACAAGAAATTCGGGAATTCCTTTAGGAATTTATCTAAACCA 362
Db 398 GTCAAGTCCAGGCAATGTTTACAAGAAATTCGGGAATTCCTTTAGGAATTTATCTAAACCA 457
QY 363 TCTGATGCTTAAGCTGAGCATAACATGAGTAAATGTTCAACCGAAGATCTCTGACAA 422
Db 458 TCTGATGCTTAAGCTGAGCATAACATGAGTAAATGTTCAACCGAAGATCTCTGACAA 517
QY 423 GTCAAGTCCAGGCAATGTTTGGGAGGATCATATAAGCCTTCCAGGAATTCGAATCT 482
Db 518 GTCAAGTCCAGGCAATGTTTGGGAGGATCATATAAGCCTTCCAGGAATTCGAATCT 577
QY 483 CTGCTTCCATTTGCAAAATGTTTCTCGGAGTACTTCAAGATTTAATCCACAA 542
Db 578 CTGCTTCCATTTGCAAAATGTTTCTCGGAGTACTTCAAGATTTAATCCACAA 637
QY 543 ATGCTTCAAGACTCTGCAAGCTCTGGATTTGATGAG----- 578
Db 638 ATGCTTCAAGACTCTGCAAGCTCTGGATTTGATGAGGAGGATCATCTGTCAAGTCTGT 697
QY 579 ----- 578
Db 698 TCACCCATTAGTCTTACTAAGCCATTCTCATTTTAAATGTCGTTGAAATCTGAATC 757
QY 579 -----GATATGTTTATACAGCTCTTTCAGAAACTAACACAGAGTATAGAG 627
Db 758 CTCAGGAGCAAGATATGTTTATACAGCTCTTTCAGAAACTAACACAGAGTATAGAG 817
QY 628 CAGCAATGAAATTCATTAATGAGTAAATGAGTAAATGAGTAAATGAGTAAATGAGTAA 687
Db 818 CAGCAATGAAATTCATTAATGAGTAAATGAGTAAATGAGTAAATGAGTAAATGAGTAA 877
QY 688 CAGCAGCTGCCAGACTTATTAATGCCAGCATGAAACCCAGGGAATGTCAGCAATCAGTTA 747
Db 878 CAGCAGCTGCCAGACTTATTAATGCCAGCATGAAACCCAGGGAATGTCAGCAATCAGTTA 937
QY 748 ACCGCAACAGAGCTGGAAGGTTCTTAAGATCTTCTAGTCTCCTCAGAGGATGCCCGC 807
Db 938 ACCGCAACAGAGCTGGAAGGTTCTTAAGATCTTCTAGTCTCCTCAGAGGATGCCCGC 997
QY 808 CACTAGGAGAAAGTGGGCTATCATTTCTGAGATGCCAATCTCAGACAGATGAGGAA 867
Db 998 CACTAGGAGAAAGTGGGCTATCATTTCTGAGATGCCAATCTCAGACAGATGAGGAA 1057
QY 868 GACCTTTCTGCTGATCTGATATATCAGCATTTGTTCAAGCTCACCTCA 915
Db 1058 GACCTTTCTGCTGATCTGATATATCAGCATTTGTTCAAGCTCACCTCA 1105

```

RESULT 3
 A1114863/c
 LOCUS
 DEFINITION HAI474 Human fetal liver cDNA library Homo sapiens EST 11-NOV-1999
 sequence.

```

ACCESSION A1114863
VERSION A1114863.1 GI:6360208
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 721)
AUTHORS Xu,Y., Zhang,C., Luo,L., Ouyang,S., Li,W., Wu,J., Zhou,S., Liu,M.
and He,F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: yut4@yahoo.com.
FEATURES
Source
1..721
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="vector: pCDNA1"
BASE COUNT 170 a 156 c 151 g 242 t 2 others
ORIGIN
Query Match 17.5%; Score 697.4; DB 9; Length 721;
Best Local Similarity 98.8%; Pred. No. 3.6e-168;
Matches 712; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 217 ATATAGATGTTTTCATGAGGAGGAGTGAAGCCAGAGGATATAGACAAATGAGGCCTA 276
Db 721 ATATAGATGTTTTCATGAGGAGGAGTGAAGCCAGAGGATATAGACAAATGAGGCCTA 662
QY 277 AGACCTTTCCCTG-CCAGTAACATATCTGTCAGTACCGGCAATGTTACAGAAATTCGG 335
Db 661 AGACCTTTCCCTGCCAGTAATTTATCTGTCAGTACCGGCAATGTTACAGAAATTCGG 602
QY 336 GAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCTAAGGCTGAGCATTAACATGAT 395
Db 601 GAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCTAAGGCTGAGCATTAACATGAT 542
QY 396 AAAATGTCACCGAAGATCTCTCGACAAGTCAAGAAATCCACCCAAATTTGGGACGATCAT 455
Db 541 AAAATGTCACCGAAGATCTCTCGACAAGTCAAGAAATCCACCCAAATTTGGGACGATCAT 482
QY 456 AAAGCCTTGCAGGAATTCGAAACTCTCTGCTTCCATTTGCAATGCAAAATTTCTTCT 515
Db 481 AAAGCCTTGCAGGAATTCGAAACTCTCTGCTTCCATTTGCAATGCAAAATTTCTTCT 422
QY 516 CGGAGTACTTCAAGAGTTAATCCACAAATGCTTCAAGACTTCAAGCTGCTGGATTTGAT 575
Db 421 CGGAGTACTTCAAGAGTTAATCCACAAATGCTTCAAGACTTCAAGCTGCTGGATTTGAT 362
QY 576 GAGGATATGGTTATACAAAGCTTCTCAGAAAATCAACACAGAGATATAGAGCAGCAAT 635
Db 361 GAGGATATGGTTATACAAAGCTTCTCAGAAAATCAACACAGAGATATAGAGCAGCAAT 302
QY 636 GAATTCATTTAGTAAATGAGTTACCAAGATCTCTCAGCAGAGCAGATGGCTGCAGCAGT 695
Db 301 GAATTCATTTAGTAAATGAGTTACCAAGATCTCTCAGCAGAGCAGATGGCTGCAGCAGT 242
QY 696 GCCAGACCTATTATGCCAGCATGAAACCCAGGGAATGTCAGCAATTCAGTAAACCGCAA 755
Db 241 GCCAGACCTATTATGCCAGCATGAAACCCAGGGAATGTCAGCAATTCAGTAAACCGCAA 182
QY 756 CAGAGCTGGAAGGTTCTTAAAGAAATCCTTAGTTCCTCAGAGGATGGCCCGCCTAGGA 815

```

Db 181 CAGAGCTGGAAGGTTCTTAAAGAACTTCTAGTTCTCTCAGAGGCGATGGCCGCCACTAGGA 122

Qy 816 GAAAGTGGCCCTATCATCTGAGAGTCCCACTCAGACAGATGTAGGAGAGCTTTG 875

Db 121 GAAAGTGGCCCTATCATCTGAGAGTCCCACTCAGACAGATGTAGGAGAGCTTTG 62

Qy 876 TCTGGATCTGGTATATCAGCATTTGTTCAAGCTCACCCTAGCAACGGAGAGAGTGAAC 935

Db 61 TCTGGATCTGGTATATCAGCATTTGTTCAAGCTCACCCTAGCAACGGAGAGAGTGAAC 2

Qy 936 C 936

Db 1 C 1

RESULT 4

AV708169 694 bp mRNA linear EST 09-OCT-2000

LOCUS AV708169 ADC Homo sapiens cDNA clone ADCALE11 5', mRNA sequence.

DEFINITION AV708169

ACCESSION AV708169

VERSION AV708169.1 GI:10725434

KEYWORDS EST

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.

TITLE Homo sapiens cDNA ADC clones

JOURNAL Unpublished (2000)

COMMENT Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China

Tel: 86-21-50801919 (ex.45)

Fax: 86-21-50801922

Email: hanzge@hgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

1. 694

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="ADCALE11"

/tissue_type="Adrenal gland"

/dev_stage="Adult"

/lab_host="SOLR"

/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 229 a 177 c 124 g 162 t 2 others

ORIGIN

Query Match: 16.3%; Score 650.8; DB 10; Length 694;

Best Local Similarity 97.8%; Pred. No. 3.5e-156;

Matches 680; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

Qy 1477 GTCATACATGGAACTATATACATAGTGTACCTGAGCTGCAACAAATTTGGCTTCAGT 1536

Db 1 GTCATACATGGAACTATATACATAGTGTACCTGAGCTGCAACAAATTTGGCTTCAGT 60

Qy 1537 CATCTTCTGCTCCAGCCAGTCATCCCGGAGCAGTGGGCGATGAATCCCTACATGGCAAC 1596

Db 61 CATCTTCTGCTCCAGCCAGTCATCCCGGAGCAGTGGGCGATGAATCCCTACATGGCAAC 120

Qy 1597 CTAACATACAGTGGAGTCAAAATCTTTTAAATACCATAGGAAATAGAGCAAGTCACT 1656

Db 121 CTAACATACAGTGGAGTCAAAATCTTTTAAATACCATAGGAAATAGAGCAAGTCACT 180

Qy 1657 CTGCTATTTCTCAGCTTCTCTCAACAGTCACTGCAATATACACAGCTCCCTATTCAAC 1716

Db 181 CTGCTAATTTCTCAGCCTTCGCTACAAAGTCACTGCAATTACACAGCTCCCTATTCAAC 240

Qy 1717 ACCCTGTGAAAAGTATGCGTCTATTAAACCCAGAGCTACAGACTGCTTTTACGACCTACAC 1776

Db 241 AGCCTGTGAAAAGTATGCGTCTATTAAACCCAGAGCTACAGACTGCTTTTACGACCTACAC 300

Qy 1777 ACCCTTCTTGGATACCAAGCAATTTCAAACTGTTCACACAGTCCCTTTTCTCTGAGGAA 1836

Db 301 ACCCTTCTTGGATACCAAGCAATTTCAAACTGTTCACACAGTCCCTTTTCTCTGAGGAA 360

Qy 1837 CGCTTCAAAATGTGACTGTGATGCCACCTGTTCGCTGAACTCCCAAACTATCAAGGACCA 1896

Db 361 CGCTTCAAAATGTGACTGTGATGCCACCTGTTCGCTGAACTCCCAAACTATCAAGGACCA 420

Qy 1897 CACCACTTACCAAAACATCTGCTGACCAAAACCCATCTGTTCTTCCCTACATACGAGTCAA 1956

Db 421 CACCACTTACCAAAACATCTGCTGACCAAAACCCATCTGTTCTTCCCTACATACGAGTCAA 480

Qy 1957 TCAGTAAGCCTAGCAAGAGGATCAGCCAAAGCTTCCCAAGAGATGAGAGTGAAGA 2016

Db 481 TCAGTAAGCCTAGCAAGAGGATCAGCCAAAGCTTCCCAAGAGATGAGAGTGAAGA 540

Qy 2017 GTTATGAAATGTGATAGTGGGATAAAGAAAGAAACAGATTACAACTTCCACTATTA 2076

Db 541 GTTATGAAATGTGATAGTGGGATAAAGAAAGAAACAGATTACAACTTCCACTATTA 599

Qy 2077 CTGTTAGGAAAACAAAGAGATGAAGCGAAGGAAATCTCGTATTCAAGTTATTCTC 2136

Db 600 CTGTTAGGAAAACAAAGAGATGAAGCGAAGGAAATCTCGTATTCAAGTTATTCTC 659

Qy 2137 CTCAGCAATTTAAATTTCTTTATGGAGCAACATGTA 2171

Db 660 CTCAGCA-TTAAATTTCTTTATGGAGCAACATGTA 693

RESULT 5

BM971435/c 657 bp mRNA linear EST 21-MAR-2002

LOCUS BM971435 UI-CF-DUI-16-0-UI.s2 UI-CF-DUI Homo sapiens cDNA clone

DEFINITION UI-CF-DUI-16-0-UI 3', mRNA sequence.

ACCESSION BM971435

VERSION BM971435.1 GI:19589022

KEYWORDS EST

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

CDNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).

The following repetitive elements were found in this cDNA sequence: 1-30, >AT-rich#low_complexity

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. 657

/organism="Homo sapiens"

FEATURES

Source

/db_xref="taxon:9606"
/clone="UI-CF-DUI-abd-1-16-0-UI"
/clone_lib="UI-CF-DUI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT 155 a 114 c 150 g 237 t 1 others
ORIGIN
Query Match 14.8%; Score 590.6; DB 14; Length 657;
Best Local Similarity 96.0%; Pred. No. 1.1e-140;
Matches 605; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1651 GTCACCTGCTAAATCTCAGCCTTCTGTACACAGTCACCTGCAATTACACAGCTCCTA 1710
Db 641 GTCACCTGCTAAATCTCAGCCTTCCGNTACACAGTCACCTGCAATTACACAGCTCCTA 582
Qy 1711 TTCACAGCCTGTAAAGATGCGTGATTAAACACAGACTACAGACTGCTTTAGCAC 1770
Db 581 TTCACAGCCTGTAAAGATGCGTGATTAAACACAGACTACAGACTGCTTTAGCAC 522
Qy 1771 CTACACACCCCTTTGGATACACAGCAATTCAACTGTTCAACCCAGCTCTTTTCCTG 1830
Db 521 CTACACACCCCTTTGGATACACAGCAATTCAACTGTTCAACCCAGCTCTTTTCCTG 462
Qy 1831 AGGAAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAAACATCAAG 1890
Db 461 AGGAAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAAACATCAAG 402
Qy 1891 GACCACAGCAGCCTTACCACCAATCTGCTGCACCAAAACCCATCTTCTCCATACG 1950
Db 401 GACCACAGCAGCCTTACCACCAATCTGCTGCACCAAAACCCATCTTCTCCATACG 342
Qy 1951 AGTCAATCAGTAAGCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATGAGAGTG 2010
Db 341 AGTCAATCAGTAAGCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATGAGAGTG 282
Qy 2011 AAAAGAGTATGAAATGTTGATAGTGGGATAAAGAAAGAAAGAAAGATTAACACTTCAC 2070
Db 281 AAAAGAGTATGAAATGTTGATAGTGGGATAAAGAAAGAAAGAAAGATTAACACTTCAC 222
Qy 2071 CTATTACTGTTAGAAAACAAAGAAAGATGAAGCGGAGGAATCTCGTATTCAAGATT 2130
Db 221 CTATTACTGTTAGAAAACAAAGAAAGATGAAGCGGAGGAATCTCGTATTCAAGATT 162
Qy 2131 ATTCTCTCAAGCAATTTAAATCTTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTC 2190
Db 161 ATTCTCTCAAGCAATTTAAATCTTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTC 102
Qy 2191 ATCAGCAGCGTCTACATCGTAAAAAACAATTTAGAGAATGAATGATCGGGTTGGATTAT 2250
Db 101 ATCAGCAGCGTCTACATCGTAAAAAACAATTTAGAGAATGAATGATCGGGTAAACCTT 42
Qy 2251 CTCAGATGCCAGGATCAATGAGAAGA 2280
Db 41 TTAATAATGTCCATTTTATATACTTAAAAAA 12

RESULT 6
LOCUS BG331628
DEFINITION 60243262BF1 NIH_MGC_18 Homo sapiens CDNA clone IMAGE:4550430 5',
mRNA sequence.
ACCESSION BG331628
VERSION BG331628.1 GI:13138066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1203)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L10M1243 row: 9 column: 07
High quality sequence stop: 724.
FEATURES
Location/Qualifiers
1..1203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550430"
/lab_host="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 330 a 272 c 281 g 320 t
ORIGIN
Query Match 14.1%; Score 560; DB 12; Length 1203;
Best Local Similarity 98.3%; Pred. No. 1e-132;
Matches 566; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 3 CTTTGGGTTGCTGGACGGACTCTGGCGCTCAGCGTCCGCCCTCAGCGCCGCGGCGC 62
Db 90 CTTTGGGTTGCTGGACGGACTCTGGCGCTCAGCGTCCGCCCTCAGCGCCGCGGCGC 149
Qy 63 TGTCAGAGCTCTGCTCTCCCTCCAGAGTTAATTTATATTTATTTATTTAAAGAAATTTAAC 122
Db 150 TGTCAGAGCTCTGCTCTCCCTCCAGAGTTAATTTATTTATTTATTTAAAGAAATTTAAC 209
Qy 123 AGTCCTGGGACTTCCCTTGAAGGATCATTTTCACCTTTTGCCTCAGAAGAAAGCTCTGGATC 182
Db 210 AGTCCTGGGACTTCCCTTGAAGGATCATTTTCACCTTTTGCCTCAGAAGAAAGCTCTGGATC 269
Qy 183 TATCAATAAAGAAGTCTCTCGTGTGGCTACATATATATATATATATATATATATATATATAT 242
Db 270 TATCAATAAAGAAGTCTCTCGTGTGGCTACATATATATATATATATATATATATATATAT 329
Qy 243 GAAAGCCAGAGGATATAGACAAATGAGGCCTTAAGACCTTTCTCCGACGTAACATATACT 302
Db 330 GAAAGCCAGAGGATATAGACAAATGAGGCCTTAAGACCTTTCTCCGACGTAACATATACT 389
Qy 303 GTCAAGTACCGGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCA 362

```

Db 390 GTGAGTACGGCGAAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTTAAACCA 449
QY 363 TCTGATGCTGCTAAGGCTGAGCATAACATGAGTAAATGTAACCCGAGAGTCCCTCGACAA 422
Db 450 TCTGATGCTGCTAAGGCTGAGCATAACATGAGTAAATGTAACCCGAGAGTCCCTCGACAA 509
QY 423 GTCAGAAATCCACCAAAATTTGGGACGATCATAAAGCCTTGCAGGAAATTCGAAACTCT 482
Db 510 GTCAGAAATCCACCAAAATTTGGGACGATCATAAAGCCTTGCAGGAAATTCGAAACTCT 569
QY 483 CTGCTTCATTTGCAATGAACAAATCTTCTCGGAGTACTTCCAGAGTTAATCCACAA 542
Db 570 CTGCTTCATTTGCAATGAACAAATCTTCTCGGAGTACTTCCAGAGTTAATCCACAA 629
QY 543 ATGCTTCAAGATGTCAGCTGCTGATTTGATGAG 578
Db 630 ATGCTTCAAGATGTCAGCTGCTGATTTGATGAG 665

RESULT 7
AA910802/c
LOCUS
DEFINITION
  AA910802 637 bp mRNA linear EST 13-APR-1998
  O14806.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
  IMAGE:1526699 3' similar to TR:Q24096 Q24096 LATS. [1]; mRNA
  sequence.
ACCESSION
  AA910802
VERSION
  AA910802.1 GI:3050092
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 637)
REFERENCE
  NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgaps-r@mail.nih.gov
  This clone is available royalty-free through LNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Seq primer: -40m13 fwd. ET from Amersham
  High quality sequence stop: 415.
  Location/Qualifiers
    1..637
      /organism="Homo sapiens"
      /db_xref="taxon:9606"
      /clone_lib="Soares_NFL_T_GBC_S1"
      /lab_host="DH10B"
      /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
      a modified polylinker; Site:1: Not 1; Site:2: Eco RI;
      Equal amounts of plasmid DNA from three normalized
      libraries (retal lung NbHL19w, testis NHR, and B-cell
      NCI-CCAP-GC81) were mixed, and ss-circles were made in
      vitro. Following HAP purification, this DNA was used as
      tracer in a subtractive hybridization reaction. The driver
      was PCR-amplified cDNAs from pools of 5,000 clones made
      from the same 3 libraries. The pools consisted of
      I.M.A.G.E. clones 297480-302087, 682632-687239,
      726408-728711, and 729096-731399. Subtraction by Bento
      Soares and M. Fatima Bonaldo."
  BASE COUNT 181 a 118 c 124 g 214 t
  ORIGIN

Query Match 14.1%; Score 559.8; DB 9; Length 637;
Best Local Similarity 95.3%; Pred. No. 9e-133;
Matches 609; Conservative 0; Mismatches 27; Indels 3; Gaps 3;

QY 3021 TGTGATTGGTGGAGTCTGGTGTATCTTTTGAATGTTGGTGGACACCTCTTTC 3080
Db 637 TGTGATTGGTGGAAATCTGG-GGAAATCTTTTGGAAATTTGTTGGAAACACCTCTTTT 579

```

```

QY 3081 TTGCAACAAACACCATTAGAAACAAATGAAGCTTTATCAACTGGCAACATCTCTTTCAC 3140
Db 578 GGTGAACACACACC-TTGGAAACCCCAAGTAGGTTTACCACCTGGCAACATCTCTTTCAC 520
QY 3141 ATTCCACCACAAGCTAAACACTCAGTCTTCTGAGCTTCTGATCTTATTTAAACTTTGCGGA 3200
Db 519 ATTCCACCACAGTCTAAACACTCAGTCTTCTGAGCTTCTGATCTTATTTAAACTTTGCGGA 460
QY 3201 GGACCCCAAGATGCTTTAGGCAAGAATGGTCTGATGAATATAAAGCTCATFCCA-TTTTT 3259
Db 459 GGACCCCAAGATGCTTTAGGCAAGAATGGTCTGATGAATATAAAGCTCATFCAATTTTT 400
QY 3260 TAAACAAATTTGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATTCCTAAAT 3319
Db 399 TAAACAAATTTGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATTCCTAAAT 340
QY 3320 CACACACCCCAACAGATACATCAAAATTTTGATCTCTGTTGATCTCTGATAAATTTATGGAGTGA 3379
Db 339 CACACACCCCAACAGATACATCAAAATTTTGATCTCTGTTGATCTCTGATAAATTTATGGAGTGA 280
QY 3380 TGATAACGAGGAAGAAATGTAATGACACTCTCAATGGATGTTATATAAATGGAAGCA 3439
Db 279 TGATAACGAGGAAGAAATGTAATGACACTCTCAATGGATGTTATATAAATGGAAGCA 220
QY 3440 TCCTGAACATGCAATCTTATGAATTTTACCTTCCGAAGGTTTTTTGATGACAATGGCTACCC 3499
Db 219 TCCTGAACATGCAATCTTATGAATTTTACCTTCCGAAGGTTTTTTGATGACAATGGCTACCC 160
QY 3500 ATATAATTTATCCGAAGCCTATTGAATATGAATCAATTAATTCACAGGCTCAGAGCAGCA 3559
Db 159 ATATAATTTATCCGAAGCCTATTGAATATGAATCAATTAATTCACAGGCTCAGAGCAGCA 100
QY 3560 GTCGGATGAAGATGATCAAAACACAGGCTCAGAGATATAAATCGCGATCTAGTATATGT 3619
Db 99 GTCGGATGAAGATGATCAAAACACAGGCTCAGAGATATAAATCGCGATCTAGTATATGT 40
QY 3620 TTAACACACTAGTAAATATAATGTAATGAGGATTTGTTAAA 3658
Db 39 TTAACACACTAGTAAATATAATGTAATGAGGATTTGTTAAA 1

RESULT 8
BE093083
LOCUS
DEFINITION
  RC5-BT0744-260400-031-G07 BT0744 Homo sapiens cDNA, mRNA sequence.
ACCESSION
  BE093083
VERSION
  BE093083.1 GI:8493535
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 578)
REFERENCE
  Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
  Nagai,M.A., da Silva,W. Jr., Zagai,M.A., Bordin,S., Costa,F.F.,
  Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
  Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
  M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
  Simpson,A.J.
  Shotgun sequencing of the human transcriptome with ORF expressed
  sequence tags.
  Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  20202663
  CONTACT: Simpson A.J.G.
  Laboratory of Cancer Genetics
  Ludwig Institute for Cancer Research
  Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
  Brazil
  Tel: +55-11-2704922
  Fax: +55-11-2707001
  Email: asimpsoneludwig.org.br
  This sequence was derived from the FAPESP/LICR Human Cancer Genome

```

Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC5-BT0744-2664-400-031-G00t3-2000-04-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 12
 High quality sequence stop: 467.
 Location/Qualifiers

FEATURES

1..578
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT0744"
 /dev_stage="Adult"

/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 163 a 120 c 129 g 165 t 1 others

BASE COUNT 163 a 120 c 129 g 165 t 1 others

Query Match 13.9%; Score 555.2; DB 10; Length 578;
 Best Local Similarity 99.3%; Pred. No. 1.3e-131;
 Matches 557; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2605 GCCTATTAAATGAGTGGGATCTTCCAGAAAGTCTGGCAGGATCTACATAGCAGAAC 2664
 DB 18 GCCTATTAAATGAGTGGGATCTTCCAGAAAGTCTGGCAGGATCTACATAGCAGAAC 77
 QY 2665 TTACTGTGCAGTTGAAGTGTTCATAAAATGGTGTTCATAGAGATATTAACCTG 2724
 DB 78 TTACTGTGCAGTTGAAGTGTTCATAAAATGGTGTTCATAGAGATATTAACCTG 137
 QY 2725 ATAAATTTGATGATCGTGAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 2784
 DB 138 ATAAATTTGATGATCGTGAATGATTAATTAATTAATTAATTAATTAATTAATTAAT 197
 QY 2785 GCTTCAGATGGACACAGATCTTAAGTACTATCAGAGTGTGACCATCCAGCCAGAGATA 2844
 DB 198 GCTTCAGATGGACACAGATCTTAAGTACTATCAGAGTGTGACCATCCAGCCAGAGATA 257
 QY 2845 GCATGATTTCAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGACAGACTGA 2904
 DB 258 GCATGATTTCAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGACAGACTGA 317
 QY 2905 AGCCATTAGCGGAGCTGCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2964
 DB 318 AGCCATTAGCGGAGCTGCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 377
 QY 2965 GGACTCCCAATATATTGACCTGAAGTGTGCTACGACAGATACACAGTCTGCTG 3024
 DB 378 GGACTCCCAATATATTGACCTGAAGTGTGCTACGACAGATACACAGTCTGCTG 437
 QY 3025 ATTGTGGAGTGTGCTGTTATTTTGAATGTTGTTGGGACAACTCTCTTCTGG 3084
 DB 438 ATTGTGGAGTGTGCTGTTATTTTGAATGTTGTTGGGACAACTCTCTTCTGG 497
 QY 3085 CACAACACCATAGAAACACAAATGAAGTATCAACTGGCAACATCTCTTCAATTC 3144
 DB 498 CACAACACCATAGAAACACAAATGAAGTATCAACTGGCAACATCTCTTCAATTC 557
 QY 3145 CACCACAAGCTAACTCAGTC 3165
 DB 558 CACCACAAGCTAACTCAGTC 578

RESULT 9
 B0309028
 LOCUS B0309028 597 bp mRNA linear EST 16-MAY-2002
 DEFINITION MR0-BT4508-060701-003-h06 BT4508 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BQ309028
 VERSION BQ309028.1 GI:20851571

KEYWORDS Homo sapiens
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 597)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-BT4508-060701-003-h06&t3=2001-07-06&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 10
 High quality sequence stop: 596.
 Location/Qualifiers

1..597
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="BT4508"
 /dev_stage="Adult"
 /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
 208 a 117 c 125 g 147 t

BASE COUNT 208 a 117 c 125 g 147 t
 ORIGIN

Query Match 13.8%; Score 548.2; DB 14; Length 597;
 Best Local Similarity 99.0%; Pred. No. 8.5e-130;
 Matches 583; Conservative 0; Mismatches 3; Indels 3; Gaps 3;
 QY 210 GCTACATATATAGATGTTTTCATGA-AGAGGAGTGAAGCCAGAGGATATAGACAAAT 268
 DB 10 GTTACATATATAGATGTTTTCATGACAGAGGAGTGAAGCCAGAGGATATAGACAAAT 69
 QY 269 GAGGCTTAAGACCTTTCCTGCCAGTAATCTACTGTCAGTCCGCAATGTTACAAGA 328
 DB 70 GAGG-CTAAGACCTTTCCTGCCAGTAATCTACTGTCAGTCCGCAATGTTACAAGA 128
 QY 329 AATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCTAAGGCTGAGCATAA 388
 DB 129 AATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCTAAGGCTGAGCATAA 188
 QY 389 CATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCAGAAATCCACCAATTTGGGAC 448
 DB 189 CATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCAGAAATCCACCAATTTGGGAC 248
 QY 449 GCATCATAAAGCTTCGAGAAATTCGAACTCTCTGCTCCATTTGCAATGAACAA 508
 DB 249 GCATTAAGGCTTCGAGAAATTCGAACTCTCTGCTCCATTTGCAATGAACAA 308
 QY 509 TTCTTCTCGAGTACTTCAGAAATTAATCCACAAATGCTTCAAGACTTGC-AAGCTGCTG 567


```

Db 309 TTCTTCTCGGAGTACTTTCAGAAAGTAAATCCACAATGCTTCAAGACTTGCAGAAAGCTGCTG 368
QY 568 GATTGTGATGAGGATGTTATACAAAGCTTTCAGAAAGTAAATCCACAAGAGTATAGAAG 627
Db 369 GATTGTGATGAGGATGTTATACAAAGCTTTCAGAAAGTAAATCCACAAGAGTATAGAAG 428
QY 628 CAGCAATGTAATTCATTAGTAAATAGTATACCAAGATCTTCAGAGAGAGAGATGCTG 687
Db 429 CAGCAATGTAATTCATTAGTAAATAGTATACCAAGATCTTCAGAGAGAGAGATGCTG 488
QY 688 CAGCAGTCCAGACCTTATTAATCCAGCATGAACAGGAGATGTCAGCAATCACTTA 747
Db 489 CAGCAGTCCAGACCTTATTAATCCAGCATGAACAGGAGATGTCAGCAATCACTTA 548
QY 748 ACCGCAAAACAGAGCTGGAAGGTTCTAAGAAATCCTTAGTTCCTCAGAG 796
Db 549 ACCGCAAAACAGAGCTGGAAGGTTCTAAGAAATCCTTAGTTCCTCAGAG 597

RESULT 10
AW502306 550 bp mRNA linear EST 01-MAR-2000
LOCUS UI-HF-BR0p-ajs-c-10-0-UI.r1 NIH_MGC_52 Homo sapiens cDNA clone
DEFINITION IMAGE:3075474 5', mRNA sequence.
ACCESSION AW502306
VERSION AW502306.1 GI:7116847
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 550)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Eco RI site shown at the beginning of the sequence.
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M3 Forward
Location/Qualifiers
1..550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3075474"
/clone_lib="NIH_MGC_52"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="MGC85"
/lab_host="DH10B (LT1)"
/notes="Vector: p773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(7.4-9.5kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT 187 a 120 c 117 g 126 t
ORIGIN
Query Match 13.6%; Score 543.2; DB 10; Length 550;
Best Local Similarity 99.5%; Pred. No. 1.6e-128;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 315 CAAATGTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCT 374
Db 3 CGAGGGTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCT 62
QY 375 AAGGCTGAGCATACATGAGTAAATGTCAACCAAGATCTTCGACAAAGTCAGAAATCCA 434

```

```

Db 63 AAGGCTGAGCATACATGAGTAAATGTCAACCAAGATCTTCGACAAAGTCAGAAATCCA 122
QY 435 CCCAAATTTGGGAGCGCATCATAAAGCTTTCGAGAAATTCGAAACTCTCTGCTTCCATTT 494
Db 123 CCCAAATTTGGGAGCGCATCATAAAGCTTTCGAGAAATTCGAAACTCTCTGCTTCCATTT 182
QY 495 GCAGAAATGAACAAATTTCTTCTCGGAGTACTTTCAGAAATTAATCCCAAAATGCTTCAAGAC 554
Db 183 GCAATATGAACAAATTTCTTCTCGGAGTACTTTCAGAAATTAATCCCAAAATGCTTCAAGAC 242
QY 555 TTGCAAGCTGCTGGATTTGATGAGGATATGGTTATACAGCTTTCAGAAACTTCAACAC 614
Db 243 TTGCAAGCTGCTGGATTTGATGAGGATATGGTTATACAGCTTTCAGAAACTTCAACAC 302
QY 615 AGAAGTATAGAACGACGCAATTTGAATTCATTAGTAAATGAGTTACCAAGATCTTCGACGA 674
Db 303 AGAAGTATAGAACGACGCAATTTGAATTCATTAGTAAATGAGTTACCAAGATCTTCGACGA 362
QY 675 GAGCAGATGGCTGCAGAGCTGCCAGACCTATTAATGCCAGATGAACACGAGGATGTG 734
Db 363 GAGCAGATGGCTGCAGAGCTGCCAGACCTATTAATGCCAGATGAACACGAGGATGTG 422
QY 735 CAGCAATCAGTTAACCCGCAACAGAGCTGGAAAGGTTCTAAAGAAATCCTTAGTTCCTCAG 794
Db 423 CAGCAATCAGTTAACCCGCAACAGAGCTGGAAAGGTTCTAAAGAAATCCTTAGTTCCTCAG 482
QY 795 AGGATGGCCGCGCAGTACGAGAAAGTGTGGCTTATCTTCTGAGAGTCCCAACTCACAG 854
Db 483 AGGATGGCCGCGCAGTACGAGAAAGTGTGGCTTATCTTCTGAGAGTCCCAACTCACAG 542
QY 855 ACAGATGT 862
Db 543 ACAGATGT 550

RESULT 11
BE568367 855 bp mRNA linear EST 15-AUG-2000
LOCUS 601341939F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3684273 5',
DEFINITION mRNA sequence..
ACCESSION BE568367
VERSION BE568367.1 GI:9812087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 855)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Inocyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM370 row: m column: 10
High quality sequence stop: 583.
Location/Qualifiers
1..855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3684273"
/clone_lib="NIH_MGC_53"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: bladder; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccctcgcc); Site_2: SfiI (ggccattatggcc
); Double-stranded cDNA was prepared from cell line RNA.
FEATURES
source

```

5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

```
BASE COUNT      277 a 217 c 184 g 177 t
ORIGIN
Query Match      13.4%; Score 533.4; DB 10; Length 855;
Best Local Similarity 96.1%; Pred. No. 6.3e-126;
Matches 612; Conservative 0; Mismatches 16; Indels 9; Gaps 6;

QY 1165 CTCCTAATCAAGCAGACGAGCATTAGTCTGTTCTCTGTTGGCAGACAAACATCATCA 1224
Db 1 CTCCTAATCAAGCAGACGAGCATTAGTCTGTTCTCTGTTGGCAGACAAACATCATCA 59

QY 1225 TGCAGAGTTCTAGCAAAATTTTCCATCAGGAGACCTGGAATGCGAATGCTACTG 1284
Db 60 TGCAGAGTTCTAGCAAAATTTTCCATCAGGAGACCTGGAATGCGAATGCTACTG 119

QY 1285 GACAACTGATTTATGATACACCAAAATGTTGTCCTGCTGGCAGTGTGAATCGGCAGC 1344
Db 120 GACAACTGATTTATGATACACCAAAATGTTGTCCTGCTGGCAGTGTGAATCGGCAGC 179

QY 1345 CACCACCTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTTCTGCTTACAAACAG 1404
Db 180 CACCACCTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTTCTGCTTACAAACAG 239

QY 1405 GGGATCTGCTCTCTCTGCTATACAAATGGAAGTATTCCTCAGTCTATGATGGTGC 1464
Db 240 GGGATCTGCTCTCTCTGCTATACAAATGGAAGTATTCCTCAGTCTATGATGGTGC 299

QY 1465 CAAACAGAAATAGTACATACAGTGAATGATATAGTCTAGTCTGACCTGCAACAA 1524
Db 300 CAAACAGAAATAGTACATACAGTGAATGATATAGTCTAGTCTGACCTGCAACAA 359

QY 1525 ATTGCGCTCAGTCATCTCTGCTCCAGCCAGTATCCCGAGCAGTGGCGATGAATCC 1584
Db 360 A-TGGCCCTCAGTCATCTCTGCTCCAGCCAGTATCCCGAGCAGTGGCGATGAATCC 418

QY 1595 CTACATGCAACCTAACATACAGTGAATGATCTTTTAAATACCCATTAGGAATA 1644
Db 419 CTACATGCAACCTAACATACAGTGAATGATCTTTTAAATACCCATTAGGAATA 478

QY 1645 GAGCAAGTCACTCTGCTAATCTCAGCTTCTGCTACACAGTCTGCTCAATACACAG 1704
Db 479 GAGCAAGTCACTCTGCTAATCTCAGCTTCTGCTACACAGTCTGCTCAATACACAG 537

QY 1705 CTCCTATTCAACAGCCTGTGAAAGTATCGGTGTTATTAACACAGCTACAGCTGCTT 1764
Db 538 CTCCTATTCAACAGCCTGTG-AAAGTATCGGTGTTATTAACACAGCTACAGCTGCTT 592

QY 1765 TAGCACCTACACACCTTCTTGGATACACAGCCCAAT 1801
Db 593 TAGCACCTACACACCTTCTTGGATACACAGCCCAAT 628

RESULT 12
LOCUS      BB557112                673 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB557112 RIKEN full-length enriched, 2 days pregnant adult female
            ovary Mus musculus cDNA clone E330026M13.3, mRNA sequence.
ACCESSION  BB557112
VERSION     BB557112.2 GI:16448760
KEYWORDS   EST.
SOURCE     Mus musculus
            house mouse.
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus
            1 (bases 1 to 673)
```

AUTHORS

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, K., Sakai, C., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

TITLE

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

JOURNAL

Unpublished (2001)

COMMENT

On Aug 1, 2000 this sequence version replaced gi:9643478.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.go.jp,

URL: http://genome.gsc.riken.go.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh

M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura

S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and

Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara

Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa

K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and

Hayashizaki, Y.

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp/) for

full details.

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Location/Qualifiers

1. 673

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="E330026M13"

/clone_lib="RIKEN full-length enriched, 2 days pregnant

adult female ovary"

/sex="female"

/tissue_type="ovary"

/dev_stage="2 days pregnant adult"

/lab_host="DH10B"

/note="Site 1: Sali; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAGCGCGCGACATCGAGTTTTTTTTTTTTTTT 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. Second strand cDNA was prepared with the

primer adapter of sequence [5'

GAGAGAGATTCCTCGATTAAATTAATTCCTCCCCCCCCCCC 3']. cDNA

was cleaved with BamHI and XhoI. Vector: a modified

pbluescript KS(+) after bulk excision from Lambda FLC I."

211 a 157 c 152 g 153 t

Query Match 13.3%; Score 530.6; DB 10; Length 673;
Best Local Similarity 86.8%; Pred. No. 3e-125;
Matches 584; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1720 CTGTGAAGATGTCGTGTTATTAACACGAGCTACAGACTGCTTTAGACCTACACACC 1779
Db 1 CCGTGAAGACATGGCGTCTGAACACGAGCTGCGAGTCTTTAGCCCAACCCATC 60

QY 1780 CTCTTGGATACACACGCAATTCAACTCTCAACCCAGTCTTTCTGAGGGAACCG 1839
Db 61 CTTCTTGAATGCCACGACCCAGTTCAGACTGTTCAAGCTACCCCTTTCTGAGGGTACAG 120

QY 1840 CTTCAATATGTACGTGATGCCACCTCTTCTGAAGCTCCAACTATCAAGGACCCACC 1899
Db 121 CTTCAAGTGTGCTGTCTATCCACCTCTTCTGAAGCTCCAACTATCAAGGACCCACC 180

QY 1900 CACCTTACCCAAAACATCTGCTGCACCAACCCATCTGTTCTCCATAGCAGTCAATCA 1959
Db 181 CGCTTATCCAAAACATCTGCTGCACCAACCCATCTGTTCTCCATAGCAGTCAATCA 240

QY 1960 GTAGCCCTAGCAAGAGGATCAGCCAGCTTGCCCAAGGAGATGAGAGTGAAGAAGTT 2019
Db 241 GTAAGCCCTGCAAGAGTGAAGAGCTAGCTTACCCAAAGGAGATGATGAGAGAGTG 300

QY 2020 ATGAAATGTTGATGAGTGGGATAAAGAAAGAACAGATTTACAACTTCCCTTACTG 2079
Db 301 CGGACAGTGGTACTCTGGGGATAAAGAAAGAACAGATTTACAACTTCCCTTACTG 360

QY 2080 TTAGAAAACCAAGAGATGAGAGGAGGATCTCGTATTCAAGTATTCTCTCTC 2139
Db 361 TTGCAAAAACCAAGAGATGAGAGGAGGATCTCGTATTCAAGTATTCTCTCTC 420

QY 2140 AAGCATTTAAATCTTTATGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGC 2199
Db 421 AGGCCCTTTAAGTCTTATGAGCAGCAGTGTAGAGAGCTCTCAAGTCTCATCAGCAGC 480

QY 2200 GTCTACATCGTAAAAACAATTAGAGATGAATGATGATGATGATGATGATGATGATG 2259
Db 481 GTCTCATCGGAAAGAGCAGCTAGAAAATGAATGATGATGATGATGATGATGATGATG 540

QY 2260 CCGAGATCAATGAGAAGATGCTTTCCAAAAGATCTAATATACATCCGCTCTTAAA 2319
Db 541 CCGAGATCAATGAGAAGATGCTTTCCAAAAGATCTAATATACATCCGCTCTTAAA 600

QY 2320 GGGCTAAAATGGACAAGTCTATGTTGTGAAGATGAAGACACTAGGATAGGACATTTG 2379
Db 601 GGGCTAAAATGGACAAGTCTATGTTGTGAAGATGAAGACACTAGGATAGGACATTTG 660

QY 2380 GTGAAGTCTGTCT 2392
Db 661 GTGAAGTCTGTCT 673

RESULT 13
BI15504
LOCUS 648 bp mRNA linear EST 05-JUL-2001
DEFINITION 602903902F1 NCI_CGAP_Mam3 Mus musculus CDNA clone IMAGE:5033425 5', mRNA sequence.

ACCESSION BI15504
VERSION BI15504
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1. (bases 1 to 648)
NIH-MGC <http://mgi.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: ccapbs-r@mail.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM11093 row: d column: 02
High quality sequence stop: 648.

FEATURES
location/Qualifiers
1. 648
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5033425"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 184 a 132 c 167 g 165 t

ORIGIN
Query Match 13.0%; Score 518.4; DB 13; Length 648;
Best Local Similarity 90.2%; Pred. No. 4.1e-122;
Matches 577; Conservative 0; Mismatches 61; Indels 2; Gaps 2;

QY 2377 TTGGTGAAGTCTCTAGCAAGAAAAGTAGATAGTAAAGCTTGTATGCAACAAAACCTC 2436
Db 11 TTGGTGAAGTCTCTAGCAAGAAAAGTAGATAGTAAAGCTTGTATGCAACAAAACCTC 59

QY 2437 TTCGAAAGAAAGATGTTCTTCTTCGAAATCAAGTCCCTCATGTTAAGGCTGAGAGAGATA 2496
Db 70 TTCGAAAGAAAGAGC-TCTGCTCCGAAATCAGGTGGCTCATGTTAAGGCGGAGAGGATA 128

QY 2497 TCCTGGCTGAAGCTGACAAATGAATGGTAGTTCGCTCTATATTTATTCATTCAGAGTAAG 2556
Db 129 TCCTAGCAGAGCGCAGCAATGAGTGGTGTGCTGCTACTACTCTTTCCAGGACAAGG 188

QY 2557 ACAATTTATCTTGTATGACTACATTCCTGGGGTGTATGATGAGCTTATTAATTA 2616
Db 189 ACAATTTGTATGTTGATGAGTACATTCCTGGGGGGGATATGATGAGCTTATTAATTA 248

QY 2617 GAATGGGCATCTTCCAGAAAGCTGGCAGCATTTCTACATAGCAGAACTTACCTGTGCAG 2676
Db 249 GAATGGGGCACTTTCTTCGAAATCTGCACGATTTCTACATAGCAGAACTTACCTGTGCAG 308

QY 2677 TTGAAAGTGTTCATAAAATGGGTTTTATTTCATAGAGATATTAAACCTGATTAATTTGA 2736
Db 309 TTGAAAGTGTTCATAAAATGGGTTTTATTTCATAGAGATATTAAACCTGATTAATTTGA 368

QY 2737 TTGATGCTGATGCTCATATTAATTTGACTGTTGGCTCTGCACCTGGCTTCAGATGGA 2796
Db 369 TTGACCGTGTGGGCATATTAATTTGACTGTTGGCTCTGCACCTGGCTTCAGATGGA 428

QY 2797 CACAGGATTTCAAGTACTATCAGAGTGGTGGTACCACCGCAAGATAGAGATGATTCA 2856
Db 429 CACATGACTCCCAAGTACTTACCAGAGTGGGATCACCCAGCAAGATAGAGATGATTCA 488

QY 2857 GTATGAATGGGGATCCCTCAAGCTGTCATGTTGGAGACAGACTGAAGCCATTAGAGC 2916
Db 489 GTAAAGCAATGGGAGATCTTCCAAATTCGTGGTGGGAGACAGACTGAAGCCATTAGAGC 548

QY 2917 GGAGAGCTGACGCCAGCAGCAGCAGCTGCTAGCACATTTCTTTGGTGGGACTCCCAATT 2976
Db 549 GGAGAGCTGCTGCCAGCAGCAGCAGCTGCTAGCCCATCTCTGTTGGGACTCCCAATT 608

```

QY 2977 ATATTGCACCTGAAGTGTGCTACGAACAGGATACACACA 3016
|||||
Db 609 ATATTGCACCTGAAGTGTGCTACGAACAGGATACACACA 648

RESULT 14
BG088986 658 bp mRNA linear EST 26-JAN-2001
LOCUS ut60q11.v1 Soares mouse NMGB.bcell Mus musculus cDNA clone
DEFINITION IMAGE:3332540 5' similar to TR:095835 095835 LARGE TUMOR SUPPRESSOR
1.. mRNA sequence.
ACCESSION BG088986
VERSION BG088986.1 GI:12571583
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 658)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MG1:1076704
Seq primer: -40RP from Gibco
High quality sequence stop: 473.
FEATURES
Location/Qualifiers
1..658
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3332540"
/lab_host="Soares mouse NMGB.bcell"
/note="Organ: germinal B-cell; Vector: p77T3b-Pac
(Pharmacia) with a modified polylinker; Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTTACCAATCTGAAGTGAGGAGCGCCCGCTGTTTTTTTTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p77T3 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 180 a 206 c 136 g 135 t 1 others
ORIGIN
Query Match 12.8%; Score 511; DB 12; Length 658;
Best Local Similarity 86.7%; Pred. No. 3.3e-120;
Matches 562; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 578 GCATATGTTATACAGCTCTTCAGAAAACCTAACACAGAGATATAGAGCAGCAATTCGA 637
|||||
Db 1 GGACATGTTATTCAAGCTCTTCAGAAAACCTAACACAGAGATATAGAGCAGCAATTCGA 60

QY 638 ATTCAATTAGTAATAGTATTACCAAGATCTTCGACAGAGAGATGGCTGCAGCAGCTGC 697
|||||
Db 61 ATTCAATTAGTAATAGTATTACCAAGATCTTCGACAGAGAGATGGCTGCAGCAGCTGC 120

QY 698 CAGACCTATTATGCGACATCAACACAGGGAATGTCGACGACATCAGTTAAACCGCAACA 757
|||||
Db 121 CAGGCGCTATTATGCGACATCAACACAGGGAATGTCGACGACATCAGTTAAACCGCAACA 180

QY 758 GAGCTGGAAGGTTCTTAAGAAATCTTAGTTCTCTCAGAGGATGGCCGCCACTAGGAGA 817
|||||
Db 181 AAGCTGGAAAGGTTCTTAAGAGTCTCTAGTTCTCTCAGAGACAGCGCCCATCTCTAGGAGA 240

QY 818 AAGTGTGCGCTATCATCTTGAGAGTCCCAACTCAGACAGATGTAGGAAGACCTTTGTC 877
|||||

```

241 AAATGTGTTTATCGTTCTGAAAGCCCAACTCACAGCGGATGTAGGAAGACCTCTGTC 300

QY 878 TGSATCTGTATATACAGCATTTCTCAAGCTCACCTTAGCAGCGACAGAGTGAACCC 937

Db 301 TGAATCCGGCATTTGCAGCATTTGCTCAAGCTCACCAAGCAATGGACAGAGTGAACCC 360

QY 938 CCACACACACCTCAAGTAAGGAGTGTACTCTCTCCACCACTCCCAAGAGGCGAGACTCC 997

Db 361 CCACACACACCTCAAGTATAGGAGTGTACTCTCTCCACCACTCCCGAGAGGCGAGACTCC 420

QY 998 CCCTCAAGAGGTACAACTCCACCTCCCTTCATCGGGAACCAAACTCTCAACAAAGCG 1057

Db 421 ACCTCCCGAGGACCACTCCCTCCCTCATGGAAACCAAGCTCTCAGACAAAGCG 480

QY 1058 CTATCTCGAAACATGGAATAGTATCTCCGAAATCTCTCTGTCTCCACCTGGGCGATG 1117

Db 481 CTATCTCGAAACATGGAATAGTATCTCCGAAATCTCTCTGTCTCCACCTGGGCGATG 540

QY 1118 GCAAGAGGCTATCTCTCCACCACTCTCAACACTTCCCTCCATGAATCTCTTAATCAAGG 1177

Db 541 GCAGGAGGCTATCTCTCCACCACTCTCAACACTTCTCTCCATGAATCTCTTAATCAAGG 600

QY 1178 ACAGAGAGCATAGTCT 1225

Db 601 TCAGAGGCGCATAGTCT 648

RESULT 15

AZ340286/c

LOCUS 1M0072J17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic

DEFINITION clone UUGC1M0072J17 F, DNA sequence.

ACCESSION AZ340286

VERSION AZ340286

KEYWORDS GSS.

SOURCE AZ340286.1 GI:10415231

ORGANISM house mouse.

Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 665)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112 USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00

Plate: 0072 row: J column: 17

Seq primer: CGTTGTAACAGCAGCGCCAGT

Class: plasmid ends

High quality sequence stop: 665.

Location/Qualifiers

1..665

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0072J17"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"

/note="Vector: PWD42hv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor-mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT	162 a	132 c	204 g	167 t
ORIGIN				
Query Match	12.5%	Score 497.6;	DB 17;	Length 665;
Best Local Similarity	86.1%	Pred. No. 9.4e-117;		
Matches 575;	Conservative 0;	Mismatches 89;	Indels 4;	Gaps 2;
Qy 1160	GAATCTCTCTAATCAAGACAGACAGAGGATTAGTTCTCTCTCTGTTGGCAGACACCA-A	1218		
Db		606		
Qy 1219	TCATCATGCAGAGTTCTAGCAAAATTTAACTTTCCATCAGGGAGACCTGGAATGAGAATG	1278		
Db		546		
Qy 1279	GTACTGGACAACTGATTTCATGATACACCAAAATTTGTCCTCTGGCAGCTGTGAATC	1338		
Db		489		
Qy 1339	GGCAGCCACCACTCCATATCTCTGACAGAGCTAATGGACAAAGCCCTCTGCTTTAC	1398		
Db		429		
Qy 1399	AAACAGGGGGATCTGCTCTCTGCTATACAAATGGAAGTATTCCTCAGTCTATGA	1458		
Db		369		
Qy 1459	TGGTGCCAAACAGAAATAGTCATACATGGAACATATATAACATTAGTACCTGGACTGC	1518		
Db		309		
Qy 1519	AAACAAATGGCCCTCAGTCATCTCTGCTCCAGCCAGTCATCCCGAGAGTGGGCATG	1578		
Db		249		
Qy 1579	AAATCCCTTACATGGCAACCTAACATACAGTGAGGTCAAAATCTTTTAAACCCATTAG	1638		
Db		189		
Qy 1639	GAATAGAGCAAGTCACTCTCTAATCTCAGCCCTCTGCTACACAGTCACTGCAATTA	1698		
Db		129		
Qy 1699	CACCACTCTTATTCACAGAGCTGTGAAAGTATGCGTGTATTAACACAGAGCTACAGA	1758		
Db		69		
Qy 1759	CTGCTTTAGCACCTTACACACCTCTTGGATACCAAGCCCAATTCAAACTGTTCAACCCA	1818		
Db		9		
Qy 1819	GTCTTTT 1826			
Db				
	8 CCCCTTTT 1			

GenCore version 5.1.3
Copyright (c) 1993 - 2003. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 5334.88 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtgcaattcaattacg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_un.*
- 28: em_vi.*
- 29: em_vt.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3213	100.0	3213	6	AR201456	AR201456 Sequence
2	3213	100.0	3213	10	AF104414	AF104414 Mus muscu
3	2483	77.3	3984	6	AR201455	AR201455 Sequence
4	2483	77.3	4307	9	AF104413	AF104413 Homo sapi
5	2472	76.9	4241	9	AF164041	AF164041 Homo sapi
6	1946.8	60.6	2442	6	E24613	E24613 warts Prote
7	1168.8	36.4	80146	9	AL583963	AL583963 Human DNA
8	1126.8	35.1	1374	6	E24614	E24614 warts Prote
9	1079.4	33.6	16490	2	AC102712	AC102712 Mus muscu
10	872.4	27.2	16490	2	AC102712	AC102712 Mus muscu
11	812.6	25.3	3155	6	AR201457	AR201457 Sequence
12	805.2	25.1	3460	10	AB023958	AB023958 Mus muscu
13	787.2	24.5	5486	6	E38226	E38226 Human tumor
14	787.2	24.5	5486	6	E38227	E38227 Human tumor
15	785.6	24.5	3412	9	AB028019	AB028019 Homo sapi
16	785.6	24.5	4098	9	AF207547	AF207547 Homo sapi
17	675.6	21.0	191604	2	AC126128	AC126128 Rattus no
18	609.8	19.0	181750	2	AC098623	AC098623 Rattus no
19	598.2	18.6	795	6	E24616	E24616 warts Prote
20	553	17.2	181750	2	AC098623	AC098623 Rattus no
21	521.2	16.2	5360	3	DR0WARTS	L39837 Drosophila
22	521.2	16.2	5720	3	DMU29608	U29608 Drosophila
23	521.2	16.2	5720	6	AR201454	AR201454 Sequence
24	468.4	14.6	582	6	E24615	E24615 warts Prote
25	353.4	11.0	183597	9	AL356285	AL356285 Human DNA
26	258.8	8.1	2422	8	AY128279	AY128279 Arabidops
27	254	7.9	236109	3	AE003775	AE003775 Drosophil
28	252.4	7.9	28133	2	AC012975	AC012975 Drosophil
29	252.4	7.9	163741	3	AC007821	AC007821 Drosophil
30	251.8	7.8	2202	8	MGPRTKINA	Z30329 M.crystalli
31	250	7.8	2355	8	SOPRTKINA	Z30330 S.oleracea
32	248.4	7.7	1935	6	AR139101	AR139101 Sequence
33	248.4	7.7	4725	9	BC028603	BC028603 Homo sapi
34	248.4	7.7	4983	6	AX056364	AX056364 Sequence
35	248.4	7.7	5181	9	AB023182	AB023182 Homo sapi
36	244.4	7.6	3018	6	A52140	A52140 Sequence 6
37	244.4	7.6	3018	6	AR084691	AR084691 Sequence
38	244.4	7.6	3018	9	HSPROKINX	Z35102 H.sapiens m
39	244.4	7.6	3593	9	BC012085	BC012085 Homo sapi
40	244	7.6	1745	3	AF275634	AF275634 Caenorhab
41	243	7.6	1738	8	AY096444	AY096444 Arabidops
42	243	7.6	2365	8	AY063969	AY063969 Arabidops
43	241.2	7.5	1322	3	E0C249683	AJ249683 Euplotes
44	241.2	7.5	1578	3	E0C249680	AJ249680 Euplotes
45	240.6	7.5	1804	3	E0C249682	AJ249682 Euplotes

ALIGNMENTS

RESULT 1
AR201456
LOCUS AR201456 3213 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6359193.
ACCESSION AR201456
VERSION AR201456.1 GI:20252344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3213)
AUTHORS Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
FEATURES Location/Qualifiers

QY	961	ATCACACCCGCTCCTATTCAACAGCCCGTGAAAAAGCATGCGCGTCTGAAAAACCAAGAGCTG	1020
DB	961	ATCACACCCGCTCCTATTCAACAGCCCGTGAAAAAGCATGCGCGTCTGAAAAACCAAGAGCTG	1020
QY	1021	CAGACTGCTTTAGCCCCAACCCATCCTTCTTGAGTGCCACAGCCAGTTTCAGACTGTTTCAG	1080
DB	1021	CAGACTGCTTTAGCCCCAACCCATCCTTCTTGAGTGCCACAGCCAGTTTCAGACTGTTTCAG	1080
QY	1081	CCTACCCCTTTTCTCAGGGTACAGCTTCAAGTGTGCGTGTCAATCCCAACCTGTTTGTCTGAA	1140
DB	1081	CCTACCCCTTTTCTCAGGGTACAGCTTCAAGTGTGCGTGTCAATCCCAACCTGTTTGTCTGAA	1140
QY	1141	GTCTCAAGCTATCAAGGTCCACACCGCTTATCCAAACATCTGCTACACACAAACCCCA	1200
DB	1141	GTCTCAAGCTATCAAGGTCCACACCGCTTATCCAAACATCTGCTACACACAAACCCCA	1200
QY	1201	TCGTGCTCCATATAGTACAGTACAGTACGCCCTGCAAGATGAACAGCTAGCTTACCTACCC	1260
DB	1201	TCGTGCTCCATATAGTACAGTACAGTACGCCCTGCAAGATGAACAGCTAGCTTACCTACCC	1260
QY	1261	AAGGAAGATGATGTCAGAAGAGTGGGACAGTGTGACTCTGGSGATAAAGAAACGAA	1320
DB	1261	AAGGAAGATGATGTCAGAAGAGTGGGACAGTGTGACTCTGGSGATAAAGAAACGAA	1320
QY	1321	CAGATTACAACCTTACCTATCACTTCTCGRAAAACAAGAAGATGAAGAACGAAGAGAG	1380
DB	1321	CAGATTACAACCTTACCTATCACTTCTCGRAAAACAAGAAGATGAAGAACGAAGAGAG	1380
QY	1381	TCTCGGATTACAGTTTACCTCCACAGGCTTTAAGTTCTTCATGGAGCAGCAGCTAGAG	1440
DB	1381	TCTCGGATTACAGTTTACCTCCACAGGCTTTAAGTTCTTCATGGAGCAGCAGCTAGAG	1440
QY	1441	AACGTCCTGAAGTCTCATCAGCAGCGCTGCTCGAAGAGCAGCTAGAAAAATGAATG	1500
DB	1441	AACGTCCTGAAGTCTCATCAGCAGCGCTGCTCGAAGAGCAGCTAGAAAAATGAATG	1500
QY	1501	ATCGGGTTGGATTATCTCAAGATCCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
DB	1501	ATCGGGTTGGATTATCTCAAGATCCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
QY	1561	GAGTCTAACTATATTCGCTTTAAAAGGCTAAAATGGACAAGTCTATGTTGTAAAGATA	1620
DB	1561	GAGTCTAACTATATTCGCTTTAAAAGGCTAAAATGGACAAGTCTATGTTGTAAAGATA	1620
QY	1621	AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGCTPAGCAAGAAAGTCGATACATAA	1680
DB	1621	AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGCTPAGCAAGAAAGTCGATACATAA	1680
QY	1681	GCCTTTGTATGCAACAAAGACTCTTCCAAAGAAAGACGTTCTGCTCCGAAATCAGGTGGCT	1740
DB	1681	GCCTTTGTATGCAACAAAGACTCTTCCAAAGAAAGACGTTCTGCTCCGAAATCAGGTGGCT	1740
QY	1741	CATGTGAAAGCGGAGAGGATATCTACAGAGCCGACAATGAGTGGTGGTCCCGCTG	1800
DB	1741	CATGTGAAAGCGGAGAGGATATCTACAGAGCCGACAATGAGTGGTGGTCCCGCTG	1800
QY	1801	TACTACTCTTCCAGSACAAGACAACCTGTACTTGTGTGAGTACATTCCTCTGGGGGG	1860
DB	1801	TACTACTCTTCCAGSACAAGACAACCTGTACTTGTGTGAGTACATTCCTCTGGGGGG	1860
QY	1861	GATATGATGAGCTATTAAATTAGAAATGGGCATCTTCTCGAAAAATCTGGCAGCATTCAC	1920
DB	1861	GATATGATGAGCTATTAAATTAGAAATGGGCATCTTCTCGAAAAATCTGGCAGCATTCAC	1920
QY	1921	ATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAATGGGTTTTATTTCATAGAGAT	1980
DB	1921	ATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAATGGGTTTTATTTCATAGAGAT	1980
QY	1981	ATTAAACCTGATAACATTTTGATTGACCGCTGATGCCATATTAATTAAGTACTGCTTGGC	2040
DB	1981	ATTAAACCTGATAACATTTTGATTGACCGCTGATGCCATATTAATTAAGTACTGCTTGGC	2040
QY	2041	TTGTGCATCTGGCTTCAGATGGACATGACCTCAAGTACTACCAAGAGTGGGGATCACCCA	2100

[illegible]

Db 3121 AATTCTTGGTACTTAAGTACTTAAAAAGAGAACCCCTGGTATCTCTTTTGTATATATAATAA 3180

QY 3181 ATAATTTTAAATCCCAAAAAAATT

Db 3181 ATAATTTTAAATCCCAAAAAAATT

RESULT 2

AF104414 3213 bp mRNA linear ROD 18-MAR-1999

LOCUS Mus musculus large tumor suppressor 1 (Lats1) mRNA, partial cds.

DEFINITION AF104414

ACCESSION AF104414

VERSION AF104414.1 GI:4324435

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.

TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase

JOURNAL Development 121 (4), 1053-1063 (1995)

MEDLINE 95262551

PUBMED 7743921

REFERENCE 2 (bases 1 to 3213)

AUTHORS Tao, W., Zhang, S., Turenchak, G.S., Stewart, R.A., St John, M.A., Chen, W. and Xu, T.

TITLE Human homologue of the Drosophila melanogaster lats tumour suppressor modulates CDC2 activity

JOURNAL Nat. Genet. 21 (2), 177-181 (1999)

MEDLINE 99140767

PUBMED 9988268

REFERENCE 3 (bases 1 to 3213)

AUTHORS St John, M.A., Tao, W., Fei, X., Fukumoto, R., Carcangiu, M.L., Brownstein, D.G., Parlow, A.F., McGrath, J. and Xu, T.

TITLE Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction

JOURNAL Nat. Genet. 21 (2), 182-186 (1999)

MEDLINE 99140768

PUBMED 9988269

REFERENCE 4 (bases 1 to 3213)

AUTHORS Tao, W. and Xu, T.

TITLE Direct Submission

JOURNAL Submitted (04-NOV-1998) Genetics, HHMI, Yale University; 295 Congress Ave BCM 254D, New Haven, CT 06536, USA

FEATURES

source Location/Qualifiers

1..3213

/organism="Mus musculus"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="newborn"

<1..3213

<1..2889

/gene="Lats1"

/gene="Lats1"

/note="putative serine threonine kinase; (MUM) Lats1"

/codon_start=1

/product="large tumor suppressor 1"

/protein_id="AADI6883.1"

/db_xref="GI:4324436"

/translation="VOHSINRKOSWKGSKESIVQRHPSLGNVYRSESPNSQADV
GPLSSGTAFAQAHPNSQQRNVNPPQVRSVTPPPRGTTPRGTTPPPSWSE
PSSQTKRISGMIEVTSIRSPVPPGAWQEGIPPLTSTPMNPSSQAQRLSSVPVGR
QLTIMQTSKFNTPRGVGGGSDFIHVONVPTGTVRTPQPPPPYPLTTPANGQS
SALQTGASAPFSFANGVPQSMVFNRSNNHMLYNIIVPGLQTAWPOSSSAQAQS
PSGHEIPIQWQNPINPVRSNNPLGSRASHANSOSPATVTATPAIQPVKMSR
VLKPELQATALATHPSWMPQVOTPTPESECTASSVPVPPVEAPSVQGPVPPY
KHLHQNPSVPVPEYSVKCKDEQSLPKEDDESEKSDGSDCKEKKQITTSPIVIR
KADQKRESRESIQTSPOAFKFFMEQHVNLKHOQRHLKRLQLENNMRVGLSQD
ANNQDEKMKLQCKESNYIRLKRAMKDSMFVKITLGIGAFGEVCLARKVDYTKALYAT
TLRKDKVILLENARVAFKEDILAEADVESVHKMGFTHRDKPNILIDRQHTKLTDFGGL
LITRMGILLENARVAFKEDILAEADVESVHKMGFTHRDKPNILIDRQHTKLTDFGGL

gene

CDS

TGFRWTHDSKYOSGDHPRODSMDPNSNCRGDRKLPKLLERRAAROHORCLAH
SLVGPNTIAEVLRLTYQLCDWNSVGVILCEMLVGOPFFLAQTPLETKMXYLIHQ
TSLHTPPQAKISPEASDLIIKLRGPDRLKNGADEIKAHFPEKTFIDFSSDLRQQA
SYIKRTHPTDTSNFDVDPDKLMSDGESEENI SDTSGWKYKHKHPEHAFYEFTRR
FFDNGYYPYKPKPEYEHISQSEQOQDEDDOHTSSDGNRLDLYV*

BASE COUNT 946 a 764 c 696 g 807 t

Query Match	100.0%	Score 3213;	DB 10;	Length 3213;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3213;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	GTGCAACATTCATTAACCGAACAACAAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCT	60		
DB 1	GTGCAACATTCATTAACCGAACAACAAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCT	60		
QY 61	CAGAGACAGGGCCATCTCTAGGAGAAAATGTGGTTTATCGTTCTGAAAGCCCAACTCA	120		
DB 61	CAGAGACAGGGCCATCTCTAGGAGAAAATGTGGTTTATCGTTCTGAAAGCCCAACTCA	120		
QY 121	CAGGCGGATGTAGGAGACCTCTGTCTGGATCCGGATTCGAGCAATTCCTCAAGCTCAC	180		
DB 121	CAGGCGGATGTAGGAGACCTCTGTCTGGATCCGGATTCGAGCAATTCCTCAAGCTCAC	180		
QY 181	CAAGCAATGGACAGAGAGTGAACCCGCCACACCACTCAAGTTAGGAGTGTACTCT	240		
DB 181	CAAGCAATGGACAGAGAGTGAACCCGCCACACCACTCAAGTTAGGAGTGTACTCT	240		
QY 241	CCACCACCTCCGAGAGGCGAGCCCACTCCCGAGGACACACTCCCGCTCCCGCTCA	300		
DB 241	CCACCACCTCCGAGAGGCGAGCCCACTCCCGAGGACACACTCCCGCTCCCGCTCA	300		
QY 301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTCTCCGA	360		
DB 301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTCTCCGA	360		
QY 361	ATCTCCCTCTTCCACTGGGGCGTGGCAGAGGGTACCTCCACCACTCTTACCCT	420		
DB 361	ATCTCCCTCTTCCACTGGGGCGTGGCAGAGGGTACCTCCACCACTCTTACCCT	420		
QY 421	TCCTCCATGAATCCCTTAGCAGGCTCAGAGGCGCATAGTCTCTCCAGTTGTAG	480		
DB 421	TCCTCCATGAATCCCTTAGCAGGCTCAGAGGCGCATAGTCTCTCCAGTTGTAG	480		
QY 481	CAACCCATCATCATGAGAGTACTAGCAAAATTTAACTTTACACAGGGGACCTGGAGTT	540		
DB 481	CAACCCATCATCATGAGAGTACTAGCAAAATTTAACTTTACACAGGGGACCTGGAGTT	540		
QY 541	CAGAATGGTGGTGTGAGTCTGATTTATCGTCACCAAAATGTCCCTGTTCTGTG	600		
DB 541	CAGAATGGTGGTGTGAGTCTGATTTATCGTCACCAAAATGTCCCTGTTCTGTG	600		
QY 601	ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGCAAGGCCCTCTGCT	660		
DB 601	ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGCAAGGCCCTCTGCT	660		
QY 661	TTACAAACAGGGGCTTCTGCTGTCTCCACCAATCATCGCCAAATGGAACGTTCTCAGTCG	720		
DB 661	TTACAAACAGGGGCTTCTGCTGTCTCCACCAATCATCGCCAAATGGAACGTTCTCAGTCG	720		
QY 721	ATGATGTGTCACACAGAGACAGTCAATACATGGAGCTTTTAAATATATGTCCTGGA	780		
DB 721	ATGATGTGTCACACAGAGACAGTCAATACATGGAGCTTTTAAATATATGTCCTGGA	780		
QY 781	CTGCAACAGCTGGGCCAGTGTCTCTGCTCTCCGAGTCACTCCCAAGCGTGGG	840		
DB 781	CTGCAACAGCTGGGCCAGTGTCTCTGCTCTCCGAGTCACTCCCAAGCGTGGG	840		
QY 841	CATGAAATTCCTACATGGCAACCTTAACATACAGTGTAGGTTCTTTTAAATACCCA	900		
DB 841	CATGAAATTCCTACATGGCAACCTTAACATACAGTGTAGGTTCTTTTAAATACCCA	900		
QY 901	TTAGGAGTAGAGCAAGTCACTCTGCTAATTTCTCAGCCTTCTGCCACTACAGTCACTGCC	960		

DB 901	TTAGGAAGTAGAGCAAGTCACTCTGCTAATTTCTCAGCTTCTTGCACCTACAGTCACTGCC	960
QY 961	ATCACACCGGCTCTTATTAACAGCCCGTGAAGCATGCGGGTCTTGAACCCAGAGCTG	1020
DB 961	ATCACACCGGCTCTTATTAACAGCCCGTGAAGCATGCGGGTCTTGAACCCAGAGCTG	1020
QY 1021	CAGACTGCTTTAGCCCAACCCATCTTCTTGGATGCCACAGCCAGTTTCAGACTGTTCCAG	1080
DB 1021	CAGACTGCTTTAGCCCAACCCATCTTCTTGGATGCCACAGCCAGTTTCAGACTGTTCCAG	1080
QY 1081	CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCTGTCATCCACCTGTTTCTGCTGAA	1140
DB 1081	CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCTGTCATCCACCTGTTTCTGCTGAA	1140
QY 1141	GCTCAAGCATATCAAGTGTCCACCCGCTTATCCAAACATCTCTACACCAAAACCCA	1200
DB 1141	GCTCAAGCATATCAAGTGTCCACCCGCTTATCCAAACATCTCTACACCAAAACCCA	1200
QY 1201	TCGTGCTCCCTCCATATAGTCAAGTAAAGTAAAGCCCTCAAGATGAACAGCTTAGCTTACC	1260
DB 1201	TCGTGCTCCCTCCATATAGTCAAGTAAAGTAAAGCCCTCAAGATGAACAGCTTAGCTTACC	1260
QY 1261	AAGGAAGATGATGAGAGAGTGGGACAGTGTGACTCTGGGGATGAAGAAAGAAA	1320
DB 1261	AAGGAAGATGATGAGAGAGTGGGACAGTGTGACTCTGGGGATGAAGAAAGAAA	1320
QY 1321	CAGATTACAACCTTCACTATCTGTCGAAAAAACAAGAGATGAAGAACGAAGAG	1380
DB 1321	CAGATTACAACCTTCACTATCTGTCGAAAAAACAAGAGATGAAGAACGAAGAG	1380
QY 1381	TCCTGGATTCAGAGTACTTCCCAAGGCTTTTAAAGTCTTCATGAGCAGCAGCTAGAG	1440
DB 1381	TCCTGGATTCAGAGTACTTCCCAAGGCTTTTAAAGTCTTCATGAGCAGCAGCTAGAG	1440
QY 1441	ACGTCCTGAAGTCTCATCAGCAGGCTGTCATCGGAAGCAGCTAGAAAATGAAATG	1500
DB 1441	ACGTCCTGAAGTCTCATCAGCAGGCTGTCATCGGAAGCAGCTAGAAAATGAAATG	1500
QY 1501	ATCGGGTGGGATTTCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
DB 1501	ATCGGGTGGGATTTCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
QY 1561	GAGTCTAATATATTCGCTTAAAGGGCTAAATGGCAAGCTATGTTCTGTAAGATA	1620
DB 1561	GAGTCTAATATATTCGCTTAAAGGGCTAAATGGCAAGCTATGTTCTGTAAGATA	1620
QY 1621	AAGACATTTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAGCTGATAATA	1680
DB 1621	AAGACATTTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAGCTGATAATA	1680
QY 1681	GCTTTGTATGCAACAAGACTCTTCGAAAGAAAGAGGTTCTGCTCCGAAATCAGGTGGCT	1740
DB 1681	GCTTTGTATGCAACAAGACTCTTCGAAAGAAAGAGGTTCTGCTCCGAAATCAGGTGGCT	1740
QY 1741	CATGTCAAAGCGGAGAGGATATCTTAGCAGAGCCGCAATGAGTGGTGGTCCGCTG	1800
DB 1741	CATGTCAAAGCGGAGAGGATATCTTAGCAGAGCCGCAATGAGTGGTGGTCCGCTG	1800
QY 1801	TACTACTCTTTCCAGGACAGGACAACTTGTACTTTGTGATGGACTACATCTCTGGGGG	1860
DB 1801	TACTACTCTTTCCAGGACAGGACAACTTGTACTTTGTGATGGACTACATCTCTGGGGG	1860
QY 1861	GATATGATGAGCCTTATTAATAGAAATGGCACTTTTCTTCCGAAATCTGCAGATTTCTAC	1920
DB 1861	GATATGATGAGCCTTATTAATAGAAATGGCACTTTTCTTCCGAAATCTGCAGATTTCTAC	1920
QY 1921	ATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGGTTTATTCATAGAGAT	1980
DB 1921	ATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGGTTTATTCATAGAGAT	1980
QY 1981	ATTAAACCTGATAACATTTGATGACCGGTGATGGCCATTAATTAATGACTGACTTGGC	2040
DB 1981	ATTAAACCTGATAACATTTGATGACCGGTGATGGCCATTAATTAATGACTGACTTGGC	2040

Db 1981 ATTAACCTGATAACATTTTGTGATGACCGTGTGATGCGCATATTAATTTGACTGACTTTGGC 2040
 QY 2041 TTGTGACTGGCTTCAGATGGACATGACTCCAAAGTACTACAGAGTGGGATCACCCA 2100
 Db 2041 TTGTGACTGGCTTCAGATGGACATGACTCCAAAGTACTACAGAGTGGGATCACCCA 2100
 QY 2101 CGGCAAGATGAGATGATTTAGTACGAATGGGAGATCCTTCCAAATTTGTCGGTGGG 2160
 Db 2101 CGGCAAGATGAGATGATTTAGTACGAATGGGAGATCCTTCCAAATTTGTCGGTGGG 2160
 QY 2161 GACAGACTGAAGCCACTGAGCGGAGAGCTGCTCGCCAGCAGCAGGATGTCAGCCAT 2220
 Db 2161 GACAGACTGAAGCCACTGAGCGGAGAGCTGCTCGCCAGCAGCAGGATGTCAGCCAT 2220
 QY 2221 TCTCTGTTGGGACCTCCCAATTAATTTGACACCTGAAGTCTACTCGGAACAGATATACA 2280
 Db 2221 TCTCTGTTGGGACCTCCCAATTAATTTGACACCTGAAGTCTACTCGGAACAGATATACA 2280
 QY 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTGTAATGTTGTAATGTTGGGCAACCT 2340
 Db 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTGTAATGTTGTAATGTTGGGCAACCT 2340
 QY 2341 CCTTTCTTGGCACAACCCCATTAAGACACAAATCAAGGTTATCATCTGGCAAACTTCT 2400
 Db 2341 CCTTTCTTGGCACAACCCCATTAAGACACAAATCAAGGTTATCATCTGGCAAACTTCT 2400
 QY 2401 CTACACATCCCTCCCTCAAGCTAAGCTGAGTCTGAGCTCTGACTCATTAATCAACATG 2460
 Db 2401 CTACACATCCCTCCCTCAAGCTAAGCTGAGTCTGAGCTCTGACTCATTAATCAACATG 2460
 QY 2461 TGTGAGGACCAAGAGACCGCTCGGCAAGAACGGTGTGATGAGATAAAGCTCATCCA 2520
 Db 2461 TGTGAGGACCAAGAGACCGCTCGGCAAGAACGGTGTGATGAGATAAAGCTCATCCA 2520
 QY 2521 TTTTAAAGACCATCGATTTCTAGTCTAGTCTGACAGCAGCAGTCTGCTTACATCCCT 2580
 Db 2521 TTTTAAAGACCATCGATTTCTAGTCTAGTCTGACAGCAGCAGTCTGCTTACATCCCT 2580
 QY 2581 AAATCAGCATCCCAACAGATACATCCCAATTTTCGACCTGTTGATCCTGATAAATTTGG 2640
 Db 2581 AAATCAGCATCCCAACAGATACATCCCAATTTTCGACCTGTTGATCCTGATAAATTTGG 2640
 QY 2641 AGCGATGCGGAGGAGGAGAAATATGATGACATCTGAGCGGATGATATAAATTTGG 2700
 Db 2641 AGCGATGCGGAGGAGGAGAAATATGATGACATCTGAGCGGATGATATAAATTTGG 2700
 QY 2701 AAGCACCCGAGCAGCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2760
 Db 2701 AAGCACCCGAGCAGCTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 2760
 QY 2761 TACCCATATAATTTCCAAAGCCTATTGATGATGATGATGATGATGATGATGATGATGATG 2820
 Db 2761 TACCCATATAATTTCCAAAGCCTATTGATGATGATGATGATGATGATGATGATGATGATG 2820
 QY 2821 CAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880
 Db 2821 CAACAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2880
 QY 2881 TATGTTTAAATAACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940
 Db 2881 TATGTTTAAATAACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 2940
 QY 2941 TTTTGAAGCTTTTGAAATATGCAAAATGTCAGAGATTTGTCGCTGCTGTGTGATACATA 3000
 Db 2941 TTTTGAAGCTTTTGAAATATGCAAAATGTCAGAGATTTGTCGCTGCTGTGTGATACATA 3000
 QY 3001 TTTTATTTTCTAGTTATGGAATTTCTTTTAAATGTTTAAATGTTTAAATTTTAAATTTTAA 3060
 Db 3001 TTTTATTTTCTAGTTATGGAATTTCTTTTAAATGTTTAAATGTTTAAATTTTAAATTTTAA 3060
 QY 3061 TTCAGTAATTTAGAAAAATTTGTTATAGGAAGTAAATTTATGACTGATGATTTATAGTC 3120
 Db 3061 TTCAGTAATTTAGAAAAATTTGTTATAGGAAGTAAATTTATGACTGATGATTTATAGTC 3120

QY 3121 AATTTCTGTACTTAAAGTACTTAAAAAGAGAGCCTGTGTATCTTTTGTATATATAATAA 3180
 Db 3121 AATTTCTGTACTTAAAGTACTTAAAAAGAGAGCCTGTGTATCTTTTGTATATATAATAA 3180
 QY 3181 ATAATTTTAAAAATCCCAAAAAA 3213
 Db 3181 ATAATTTTAAAAATCCCAAAAAA 3213

RESULT 3
 AR201455 LOCUS AR201455 Sequence 3 from patent US 6359193. DNA linear PAT 20-APR-2002
 DEFINITION AR201455
 VERSION AR201455.1 GI:20252343
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1. (bases 1 to 3984)
 AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
 TITLE Nucleotide sequences of lats genes
 JOURNAL Patent: US 6359193-A 3 19-MAR-2002;
 FEATURES Location/Qualifiers
 1. 3984
 Source /organism="unknown"
 BASE COUNT 1280 a 847 c 798 g 1059 t
 ORIGIN

Query Match 77.3%; Score 2483; DB 6; Length 3984;
 Best Local Similarity 87.4%; Pred. No. 0;
 Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;

QY 1 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTCTAGTTCCT 60
 Db 732 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTCTAGTTCCT 791

QY 61 CAGAGACAGCGGCCATCTCTAGGAGAAATGTTTATCTGTTCTGAAAGCCCACTCA 120
 Db 792 CAGAGACAGCGGCCATCTCTAGGAGAAATGTTTATCTGTTCTGAAAGTCTTCTGAGAGTCCCACCTCA 851

QY 121 CAGCGGATGTAGGAGACCTCTGTCTGGATCCGGCATTTGCGCATTTGCTCAAGCTCAC 180
 Db 852 CAGAGATGTAGGAGACCTTTGTCTGGATCTGTTATATGAGCATTTGTTCAAGCTCAC 911

QY 181 CCAACATGAGACAGAGTGAACCCCAACACCTCAAGTTAGGAGTGTACTCT 240
 Db 912 CCAACATGAGACAGAGTGAACCCCAACACCTCAAGTTAGGAGTGTACTCT 971

QY 241 CCACCACCTCCGAGAGGCGAGACCCCTCCCGAGGAGGAGTCCCTCCCTCCCTCCCTCA 300
 Db 972 CCACCACCTCCGAGAGGCGAGACTCTCCCTCCGAGGAGGAGTCCCTCCCTCCCTCCCTCA 1031

QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGATATCTCCGA 360
 Db 1032 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGATATCTCCGA 1091

QY 361 ATCTCCCTCTTCCACTGCGGCGTGGAGAGGAGTCCCTCCCTCCCTCCCTCCCTCA 420
 Db 1092 ATCTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 1151

QY 421 TCTCCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 480
 Db 1152 TCCCCCATGATCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCA 1211

QY 481 CAACCATCATCATGAGTACTAGCAAAATTTAACTTTTACACAGGAGGAGTCCCTCCCTCCCTCCCTCA 540
 Db 1212 CAACCATCATCATGAGTACTAGCAAAATTTAACTTTTCCATCAGGAGGAGTCCCTCCCTCCCTCCCTCA 1271

QY 541 CAGAAATGTTGGTCTGCTGATTTTATCGTGACCAAAA ---TGTCCCTCCCTCCCTCCCTCCCTCA 597
 Db 1272 CAGAAATGTTGGTCTGCTGATTTTATCGTGACCAAAAATTTGTTCTCCCTCCCTCCCTCCCTCA 1331

Qy 598 GTGACTCGGACGCCACACCTCCATATATCTCTGACCCAGCTAAATGGACAAAGGCCCTCT 657
Dy 1332 GTGAATCGGACGCCACACCTCCATATATCTCTGACAGCAGCTAAATGGACAAAGGCCCTCT 1391
Qy 658 GCTTTACAAACAGAGGGCTTCTGCTCCACCATCATTCGCCCAATGGAAAGCTTCCCTCAG 717
Dy 1392 GCTTTACAAACAGAGGGCTTCTGCTCCATATATATATATATATATATATATATATATAT 1451
Qy 718 TCGATGATGGTGGCCCAACAGGAACAGTATACATGAGGCTTTTAAATATATATATATATAT 777
Dy 1452 TCTATGATGGTGGCCCAACAGGAATAGTCTATACATGGAACCTATATATATATATATAT 1511
Qy 778 GGACTGCAAAACAGCGCTGGCCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 837
Dy 1512 GGACTGCAAAACAAATTTGGCCCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1571
Qy 838 GGGCATGAATTCCTACATGCGCAACCTAACATACCATGAGGTCAATCTTTTAAATAC 897
Dy 1572 GGGCATGAATTCCTACATGCGCAACCTAACATACCATGAGGTCAATCTTTTAAATAC 1631
Qy 898 CCATTAGGAAGTATAGCAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957
Dy 1632 CCATTAGGAATATAGCAAGTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1691
Qy 958 GCCATCACACCCGCT 1017
Dy 1692 GCAATTAACACAGCT 1751
Qy 1018 CTGACAGCTGCTTTAGCCCGCAACCCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1077
Dy 1752 CTACAGAGCTGCTTTAGCACCTTACACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1811
Qy 1078 CAGCCTACCCCTTTCTGAGGATACAGTCTCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCT 1137
Dy 1812 CAACCCAGTCTCTTTCTGAGGATACAGTCTCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCT 1871
Qy 1138 GAAGCTCCAGCTATCAAGGTCACACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1197
Dy 1872 GAAGCTCCAACTATCAAGGATACACACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1931
Qy 1198 CCATCTGCTCTCTATAGTCAAGTATAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1257
Dy 1932 CCATCTGCT 1991
Qy 1258 CCCAAGGAAGATGATGATGAGGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1317
Dy 1992 CCCAAGGAAGATGAGTCAAGGATGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 2051
Qy 1318 AAACAGATTACACTTCACT 1377
Dy 2052 AAACAGATTACAACTTCACT 2111
Qy 1378 GAGTCTCGGATTCAGATTCT 1437
Dy 2112 GAATCTGCTATTAAGATTCT 2171
Qy 1438 GAGAACGCTCTCAAGTCTCACTCAGCAGCTCTGATCGGAAGAGAGCAGCTAGAAAATGAA 1497
Dy 2172 GAATGCTCACTCAATCTCATCAGCAGCTCTGATCGTAAGAAACAACTATGAGAAATGAA 2231
Qy 1498 ATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAG 1557
Dy 2232 ATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAA 2291
Qy 1558 AAAGAGTCACTATATCT 1617
Dy 2292 AAAGAACTTAATATACATCT 2351
Qy 1618 ATAAGACATTAGGAATAGGAGCTTTGGTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1677
Dy 2352 ATAAGACATTAGGAATAGGAGCTTTGGTGAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2411

Qy 1678 AAAGCTTTGTATGCAACAAAGACTCTTCGAAGAAAGAGCTTCTGCTCCGAAATCAGGTG 1737
Dy 2412 AAGCTTTGTATGCAACAAAGACTCTTCGAAGAAAGAGCTTCTTCTTCGAAATCAAATTC 2471
Qy 1738 GCTCATGTGAAGCGGAGAGGATATCTTACGAGAGCGGACAAATGAGTGGTGGTCTCGC 1797
Dy 2472 GCTCATGTGAAGCGTGAAGAGATATCTTGGTGAAGTGAAGTGAAGTGGTGGTCTCGT 2531
Qy 1798 CTGTACTACTCTTCCAGGACAAAGCAACTTGTACTTTGTGTGATGGACTACATCTCTGGG 1857
Dy 2532 CTATATATATCTTCCAAAGATGAAGCAATTTATATCTTTGTAATGGACTACATCTCTGGG 2591
Qy 1858 GGGATATGATGAGCTTATTAATAGAAATGGGCACTTTCTCTGAAATCTGGCAGATTC 1917
Dy 2592 GGTATATGATGAGCTTATTAATAGAAATGGGCACTTTCTCCAGAAAGCTGGCAGATTC 2651
Qy 1918 TACATAGCAAGCACTTACCTGTCAGTCTCAAGTCTTCAAAATGGGTTTATTCATAGA 1977
Dy 2652 TACATAGCAAGCACTTACCTGTCAGTCTTCAAAATGGGTTTATTCATAGA 2711
Qy 1978 GATATTAACCTGATTAACATTTTGTATGACCTGATGGCCATATTAATTAATTAATTAATTA 2037
Dy 2712 GATATTAACCTGATTAATTTTGTATGATGCTGATGGTCAATTAATTAATTAATTAATTA 2771
Qy 2038 GGTGTGCTCACTGGCTTCAGATGAGACATGACTCCAAAGTACTACGAGTGGGATCAC 2097
Dy 2772 GGCCTCTGCACTGGCTTCAGATGAGACATGACTCCAAAGTACTACGAGTGGGATCAC 2831
Qy 2098 CCAGGCAAGATAGCATGATTTTCAATTAACAAATGGGAGATCTTCCAAATTTGGGTTG 2157
Dy 2832 CCAGGCAAGATAGCATGATTTTCAATTAACAAATGGGAGATCTTCCAAATTTGGGTTG 2891
Qy 2158 GGGGACAGACTGAAGCCACTGGAGGAGAGCTGCTCCAGCAGCAGCAGGATGTAGGCC 2217
Dy 2892 GGAGACAGACTGAAGCCACTGAAGGAGAGCTGCTCCAGCAGCAGCAGGATGTAGGCC 2951
Qy 2218 CATCTCTGTTGGGACTCCCAATATATGCACTGAAGTGTACTCTGCAACAGGATAT 2277
Dy 2952 CATCTCTGTTGGGACTCCCAATATATGCACTGAAGTGTACTCTGCAACAGGATAT 3011
Qy 2278 ACACAGCTGTGACTGTGAGTGTGTTGTTTCTTCTGAAATCTGTTGGGACAA 2337
Dy 3012 ACACAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3071
Qy 2338 CCTCTCTTTTGGGCAACAAACCCCTTATAGAAACAAATGAAGTTATCATCTGGCAAACT 2397
Dy 3072 CCTCTCTTTTGGGCAACAAACCCCTTATAGAAACAAATGAAGTTATCATCTGGCAAACT 3131
Qy 2398 TCTCTACACATCTCTCTCAAGCTAAGCTGAGTCTGAGGCTCTGAGCTCTCATTTACAA 2457
Dy 3132 TCTCTTCAATCTCCACCACAAAGCTAAGCTGAGTCTGAGGCTCTGAGCTCTCATTTAAA 3191
Qy 2458 CTGTGTCAGGACAGAGAGCCGCTCGCAAGAGGCTGCTGATGAGATGAAGGCTCAT 2517
Dy 3192 CTGTGTCAGGACAGAGAGCCGCTCGCAAGAGGCTGCTGATGAGATGAAGGCTCAT 3251
Qy 2518 CCATTTTAAAGACCATCGATTTCTTCTAGTGTCTGAGCAGCAGCTCTGCTTCTATACATC 2577
Dy 3252 CCATTTTAAAGACCATCGATTTCTTCTAGTGTCTGAGCAGCAGCTCTGCTTCTATACAT 3311
Qy 2578 CCTAAATCAGGATCCAAACAGATACATCCAAATTTGACCCCTGTTGATCTCTGATTAATG 2637
Dy 3312 CCTAAATCAGGATCCAAACAGATACATCCAAATTTGATCTCTGTTGATCTCTGATTAATA 3371
Qy 2638 TGGAGGATGCGAGGAGGAGAAATATAGTCACTCTGAGGAGTGTGATTAATAAT 2697
Dy 3372 TGGAGTGAATGATGAGGAGGAGAAATATAGTCACTCTCAATGAGTGTGATTAATAAT 3431
Qy 2698 GGGAGGACCCCGGAGCAGCTTTCTATGATTTCACTTTTCGGAGGTTTTTTCATGACAAAT 2757
Dy 3432 GGAAGGATCTGGAACATGATCTCTGAAATTTTACCTTTCCGAGGTTTTTTCATGACAAAT 3491
Qy 2758 GGCTACCCATATATATTCACAAAGCCTATTGATGATGATGATGATGATGATGATGATGATG 2817

```

Db 3492 GGCTACCATATAATATCCGAAGCCTATGAATATGAATACATTAATTCACAAAGCTCA 3551
QY 2818 GAACACAGCTGTGATGAGATGATCAACACACAGCTCCGATGGAAACACCGAGATCTA 2877
Db 3552 GAGCAGCAGCTGGATGAGATGATCAAAACACAGCTCAGAGATTAATAATTCGCGATCTA 3611
QY 2878 GTGTATGTTTAAATAACCTAGGAGATCATTTGTA-----GAAATTGCAAGAGGCTCAAGTG 2933
Db 3612 GTATATGTTTAAACACCTAGTAATAATGTAATGAGATTTGTAAGAGGCTCAAGATG 3671
QY 2934 CAGGGGTTTTCAGAGTTTGTAG-----AAATATGCAATGATGACAGAG-----TTTGTGT 2985
Db 3672 CCAGGTGTTTTCAGGTTCTGAGAGTAAATATGCAAAATATGACAGAGTATATATGTT 3731
QY 2986 GCTCTGTGTACAAATATTTATTTCTTAAAGTATGCGAAATGTTTAAATGTTAAATTT 3045
Db 3732 GCTCTGTGTACAAATATTTATTTCTTAAATGTTTAAATGTTTAAATGTTAAATTT 3791
QY 3046 ATTCACACCTTTTAAATTCAGTAATTTAGAAAAAATTTGTAAGGAAAGTAAATTTATGAA 3105
Db 3792 ATTCACACCTTTTAAATTCAGTAATTTAGAAAAAATTTGTAAGGAAAGTAAATTTATGAA 3851
QY 3106 CTGAGTATATAGTCAATTTCTGTTGTTTAAAGTACTTTAAAGAGAGAG 3154
Db 3852 CTGAATATATAGTCAATTTCTGTTGTTTAAAGTACTTTAAAGAGAGAGTAG 3900

```

```

RESULT 4
AF104413 AF104413 4307 bp mRNA linear PRI 18-MAR-1999
LOCUS Homo sapiens large tumor suppressor 1 (LATS1) mRNA, complete cds.
ACCESSION AF104413
VERSION AF104413.1 GI:4324433
KEYWORDS
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Xu,T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.
TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila
JOURNAL lats gene encodes a putative protein kinase
MEDLINE Development 121 (4), 1053-1063 (1995)
REFERENCE 95262551
AUTHORS Tao, W., Zhang, S., Turenchalk, G.S., Stewart, R.A., St John, M.A.,
TITLE Human homologue of the Drosophila melanogaster lats tumour
JOURNAL suppressor modulates CDC2 activity
MEDLINE Nat. Genet. 21 (2), 177-181 (1999)
REFERENCE 99140767
AUTHORS 9988268
TITLE 3 (bases 1 to 4307)
JOURNAL St John, M.A., Tao, W., Fei, X., Fukumoto, R., Carcangiu, M.L.,
MEDLINE Brownstein, D.G., Parlow, A.F., McGrath, J. and Xu, T.
REFERENCE Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian
AUTHORS tumours and pituitary dysfunction
TITLE Nat. Genet. 21 (2), 182-186 (1999)
JOURNAL 99140768
MEDLINE 9988269
REFERENCE 4 (bases 1 to 4307)
AUTHORS Tao, W. and Xu, T.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1998) Genetics, HHMI, Yale University, 295
MEDLINE Congress Ave BCM 254D, New Haven, CT 06536, USA
REFERENCE Location/Qualifiers
AUTHORS 1. 4307
FEATURES
SOURCE /organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="brain"
/dev_stage="fetus"

```

```

gene
CDS
/ gene="LATS1"
/ gene="LATS1"
/ note="putative serine threonine kinase"
/ product="large tumor suppressor 1"
/ protein_id="AADI6882.1"
/ db_xref="GI:4324433"
/ translation="MKSRPEGYROMRPKTFPASNYTVSSROMLOETRESLRNLKSP
SDAAKHNKMSKSTEDPQVNPFPKFGTHHKALQELINSLPFPANETNSRSSTSEVN
PQMLQDLQAAGFDEDMVIALQKLTNNRSIEAIEFISKMSYDPPRRROMAAARLPIN
ASMKPQNVQSVNRKQSKGSLVPLQHGPPGLGESVAYHSESPNDQVRLSGS
YSGFVQAHPSNGORVNPVPPPPVRSVTPPPPRGTPPPPPPPPPPPPPPPPPPPPP
YSGNMEVVISRI SPVPPGMOEGVPPPLNTSPNPNNOGORGISVYVGPVGRPIIMOS
SSKFNFSRPGMGNGTQDFTMIHQNVVAGTVNRQPPVPLTAAAGQSPSALOTG
GSAAPSSITGSIPOSMVPMNRHSHNMLYNSVPGIQTNPQSSAPSPSSGHE
IPTWPNLIPVRSNPLNRSASHNSATVTAITPAIPQVPMKSRMVLKPEL
QTALAPTHPSWIPQIOTVQSPPEPTASNTVMPVVAEPNCGPPPTPKHLHQ
NPSVPPYESIKSPKEDQSLPKDESEKSYENDSGDKKQITTSPTITVKNKQDE
ERRSRIOQSPQAFKPEMEOHVNLKSHOORLHRKOLENEMRVGLSDAODOMR
KMLCKESNYIRLRAKMDKSMFKVILKTLGICAFGEVCLARKVDTKALYAKTLRKD
VLLRQVAHVAERDILAEADNEWVRILYISFQKNDLIFVMDITPGDMSLIRMG
IFPESLARFYIAELTCAVESVHKMGFTHRIDKPNLILDRDGHKLKLDGLCTGRWT
HDSKYQSGDHPRODSMDFSNWDGPPSCDRKLPERRAARQHORCLASLVGTP
NYIAPEVLLRTGYTQCDWMSGVILFEMLVGPFPFIAOTPLETQMKVINNOTSLHIP
POAKLSPASDLIIKLCRGPEDRLGKNGADEIKAHFPFKTIDFSSDLRQOQSASVLPKI
THPTDTSNFDVPDQKLSDDNEENYNDTLNGYKNGKHPEHAFYEFTFRFFDDNG
IYINYPKPIEYIINSOGSEQSDDEDDONTGEIKNRDLVTV"
BASE COUNT 1337 a 963 c 916 g 1091 t
ORIGIN
Query Match 77.3%; Score 2483; DB 9; Length 4307;
Best Local Similarity 87.4%; Pred: No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
QY 1 GTGCAACATTCATTAACCGAAACAAAGAGTCTTAAAGAGTCTCTAGTCTCT 60
Db 1050 GTGCAACATTCATTAACCGAAACAAAGAGTCTTAAAGAGTCTCTAGTCTCT 1109
QY 61 CAGACACAGCGCCCATCTCTAGGAGAAAATGTGTTTATCGTCTGAAAGCCCCAACATCA 120
Db 1110 CAGAGGCATGCGCCGCCCATCTAGGAGAAAATGTGTTTATCGTCTGAAAGCCCCAACATCA 1169
QY 121 CAGCGGATGTAGGAAGACCTCTGTCTGGATCCGCAATTCAGCATTTGCTCAAGCTCAC 180
Db 1170 CAGACAGATGTAGGAAGACCTTTGTCTGGATCTGTATATCAGCATTTGTTCAAGCTCAC 1229
QY 181 CCAAGCAATGGACAGAGTGAACCCCGCCACACACCTCAAGTTAGGAGTGTACTCT 240
Db 1230 CCTAGCAACGGACAGAGTGAACCCCGCCACACCTCAAGTTAGGAGTGTACTCT 1289
QY 241 CCACACCTCCGAGAGGCGACCCCGCCACACCTCCCGAGGCGACCTCCCGCCCGCTCA 300
Db 1290 CCACACCTCCGAGAGGCGACCTCCCGCCACACCTCCCGAGGCGACCTCCCGCCCGCTCA 1349
QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAAACATGAGTACGTAATCTCCCGA 360
Db 1350 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAAACATGAGTACGTAATCTCCCGA 1409
QY 361 ATCTCCCTGTTCACCTGGGCGGTGGCAGGAGGGTACCCCTCCACCACTCTTACCCT 420
Db 1410 ATCTCTCTCTCCACCTGGGCGGTGGCAGGAGGGTACCCCTCCACCACTCTTACCCT 1469
QY 421 TCTCCCATGAATCCCGCTAGCAGGCTCAGAGGGCCATAGTCTGTTCAGCTTGGTAGA 480
Db 1470 TCCCGCATGAATCCCGCTAGCAGGCTCAGAGGGCCATAGTCTGTTCAGCTTGGTAGA 1529
QY 481 CAACCCATCATCATGAGAGTACTAGCAAAATTTAACTTTACACAGGGGCGACTGGAGTT 540
Db 1530 CAACCAATCATCATGAGAGTCTAGCAAAATTTAACTTTACACAGGGGCGACTGGAGTT 1589
QY 541 CAGAATGGTGGTGGTCTGATTTTATCGTGCACCAAAA---TGTCCCGACTGGTCT 597

```

Db 1590 CAGAAATGGTACTGCAAACTGATTTATGATACACCAAAATGTTGCTCCCTGCTGGCACT 1649
QY 598 GTGACTCGGAGCAGCACCTCCATATCCCTGACCCAGCTAATGAGCAAGCCCTCT 657
Db 1650 GTGAATCGGAGCAGCACCTCCATATCCCTGACAGCAGCTAATGAGCAAGCCCTCT 1709
QY 658 GCTTTACAAACAGGGGCTTCTGCTGCTCCAGCATCATTCGCAATGGAAGGTTCCCTCAG 717
Db 1710 GCTTTACAAACAGGGGCTGCTGCTGCTCCATATACAAATGGAAGTATTCCTCAG 1769
QY 718 TCGATGATGGTGGCCCAAGAGCAAGCTATACATGAGGCTTTTATAATTAATGTCCT 777
Db 1770 TCTATGATGGTGGCCCAAGAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829
QY 778 GGACTGCAACAGCCTGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
Db 1830 GGACTGCAACAGCCTGGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1889
QY 838 GGGCATGAAATTCCTACATGGCAACCTAACATACAGTGAAGTCAAAATTCCTTTAATAAC 897
Db 1890 GGGCATGAAATTCCTACATGGCAACCTAACATACAGTGAAGTCAAAATTCCTTTAATAAC 1949
QY 898 CCATTAGGAAGTGAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
Db 1950 CCATTAGGAAGTGAAGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2009
QY 958 GCCATCACCCGCTCTTATTCACAGCCGCTGAAAGCAGTCCGCTGCTGCTGCTGCTGCT 1017
Db 2010 GCAATTCACAGCCTCTTATTCACAGCCGCTGAAAGTATGCTGCTGCTGCTGCTGCTGCT 2069
QY 1018 CTGACAGTCTGCTTGGCCCAACCCCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
Db 2070 CTACAGACTGCTTTAGCACCTACACACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2129
QY 1078 CAGCTTACCCCTCTTCTGAGGCTACAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Db 2130 CAACCCAGTCTCTTCTGAGGAAACGCTTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 2189
QY 1138 GAAGCTCCAGCTATCAAGTTCACACCCGCTTATCCAAACATCTGCTACACCAAAAC 1197
Db 2190 GAAGCTCCAACTATCAAGGACCAACCCGCTTACCCAAACATCTGCTGCTGCTGCTGCTGCT 2249
QY 1198 CCATCTGCTCCCTCATATGAGTCAAGTAAAGCCCTGCAAGTGAAGCAAGCTGCTGCT 1257
Db 2250 CCATCTGCTCCCTCATATGAGTCAAGTAAAGCCCTGCAAGTGAAGCAAGCTGCTGCT 2309
QY 1258 CCAAGGAAGTATGAGTGAAGAGTGGGACAGTGGTGAAGTGGGATTAAGAAAG 1317
Db 2310 CCAAGGAAGTATGAGTGAAGAGTGGGACAGTGGTGAAGTGGGATTAAGAAAG 2369
QY 1318 AACAGATTACAACTTCACTTCACTGCTTTCGGAAGAAAGTGAAGAAAGCAAGGA 1377
Db 2370 AACAGATTACAACTTCACTTCACTGCTTTCGGAAGAAAGTGAAGAAAGCAAGGA 2429
QY 1378 GAGTCTCGGATTCAGAGTCTACCCAGAGCTTTTAAAGTCTTCTGAGGAGCAGCAGTA 1437
Db 2430 GAATCTCGTATTCAGAGTATTCCTCAAGCATTAAATTCCTTATGAGGCAACATGTA 2489
QY 1438 GAGAACTCTGAGTCTCATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497
Db 2490 GAAATGTACTCAATCTCATCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549
QY 1498 ATGATGGGGTGGATTCCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAG 1557
Db 2550 ATGATGGGGTGGATTCCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAG 2609
QY 1558 AAAGAGTCTAACTATATTCGCTTTAAAGGGCTTAAAGTGAAGCAAGTCTATGTTTGAAG 1617
Db 2610 AAAGAACTTAATACATCCGCTTTAAAGGGCTTAAAGTGAAGCAAGTCTATGTTTGAAG 2669
QY 1618 ATAAAGACATTAAGAAATGAGGCTTTTGGTGAAGTCTGCTGAGCAAGAAAGTCCATCT 1677
|||||

Db 2670 ATAAAGACACTAGAAATAGGAGCAATTTGGTGAAGTCTGCTAGCAAGAAAGTAGATACT 2729
QY 1678 AAAGCTTTGTATGCAACAAAGACTCTTGAAGAAAGAGCTTCTGCTCCGAAATCAGGTG 1737
Db 2730 AAGCTTTGTATGCAACAAAGACTCTTGAAGAAAGAGATGTTCTTCTCGAAATCAAGTC 2789
QY 1738 GCTCATGTGAAGCGGAGAGGGATATCCTAGCAAGCCGCAATGAGTGGTGGTCCGC 1797
Db 2790 GCTCATGTGAAGCGTGAAGAGATATCCTGCTGAAGCTGACATGAATGGTAGTTCGT 2849
QY 1798 CTGTACTACTCTTCCAGGACAGCAAACTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1857
Db 2850 CTATATTATTCATTCCAAGATAAGCAAAATTTACTTTGTAATGGACTACATATCTCTGG 2909
QY 1858 GGGATATGATGAGCCTTAAATTAGAATGGCATCTTCTGCAAAATCTCGCACGATTC 1917
Db 2910 GGTGATATGATGAGCCTTAAATTAGAATGGCATCTTCCAGAAAGTCTGGCAGATTC 2969
QY 1918 TACATAGCAGAACTTACCTGTGCAAGTTGAAAGTCTTCAATAAATGGTCTTATTCATAGA 1977
Db 2970 TACATAGCAGAACTTACCTGTGCAAGTTGAAAGTCTTCAATAAATGGTCTTATTCATAGA 3029
QY 1978 GATATTAACTGATACATTTTGAATGACGCTGATGCCATATTAATTTGACTGACTTT 2037
Db 3030 GATATTAACTGATACATTTTGAATGACGCTGATGCCATATTAATTTGACTGACTTT 3089
QY 2038 GGGTGTGCACTGCTTACAGATGACACATGACTCCAGTACTACAGAGTGGGGATCAC 2097
Db 3090 GGCCTCTGCACTGCTTACAGATGACACATGACTTAAAGTACTATCAGAGTGGTGACAT 3149
QY 2098 CCACGGCAAGATAGATGATTTTCAAGTAAAGTGGGAGATCTTCCAAATTTGCTGGTGT 2157
Db 3150 CCACGGCAAGATAGATGATTTTCAAGTAAAGTGGGAGATCTTCCAAATTTGCTGGTGT 3209
QY 2158 GGGACAGATGAAGCAGCTGGAGCGGAGAGCTGCTGCCAGCAGCAGGATGCTAGCC 2217
Db 3210 GGAGACAGATGAAGCAGCTGGAGCGGAGAGCTGCTGCCAGCAGCAGGATGCTAGCA 3269
QY 2218 CATCTCTGCTGGTGGACTCCCAATTTATTCGACCTGAAGTGTACTCGCAACAGGATAT 2277
Db 3270 CATCTCTGCTGGTGGACTCCCAATTTATTCGACCTGAAGTGTACTCGCAACAGGATAT 3329
QY 2278 ACACAGCTGTGTGACTGGTGGAGTGTGGTGTATTTCTTTGTGAATTTGGTGGGACAA 2337
Db 3330 ACACAGTGTGTGATGGTGGAGTGTGGTGTATTTCTTTTGAATTTGGTGGGACAA 3389
QY 2338 CCTCTTCTTGGCACAAACCCCAATTAGAACACAAATGAGGTTATCATCTGCAAACT 2397
Db 3390 CCTCTTCTTGGCACAAACCCCAATTAGAACACAAATGAGGTTATCATCTGCAAACT 3449
QY 2398 TCTTACACATCCCTCCTCAAGCTAAGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2457
Db 3450 TCTTCTCACATTCACCCACAGCTAACTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3509
QY 2458 CTGTGTCGAGACACAGAGCAGCCTCGGCAAGAACCGGTGCTGATGAGATAAAGGCTCAT 2517
Db 3510 CTTTCCGCGAGCAGCAGGAGTGGCTTAGGCAAGAAATGGTGTGATGAATAAAGGCTCAT 3569
QY 2518 CCATTTTGAAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2577
Db 3570 CCATTTTGAAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3629
QY 2578 CCTAAATTCAGCATCCCAAGATACATCCAAATTTCCAGCCCTGTTGATCTCTGATAAATG 2637
Db 3630 CCTAAATTCAGCATCCCAAGATACATCCAAATTTGATCTCTGCTGCTGCTGCTGCTGCTGCT 3689
QY 2638 TGGAGCATGCGAGGAGGAGGAAATATCAGTGACACTCTGAGCGGATGGTATFAAAT 2697
Db 3690 TGGAGTATGATACAGGAGGAGGAAATGTAATGACACTCTCAATGGATGGTATFAAAT 3749
QY 2698 GGGAGCAGCCCGAGCAGCTTCTATGAGTTCACCTTTCCGAGGTTTCTGATGACAAT 2757
Db 3750 GGAAGCATCTTCAACATGATCTATGATTTTACCTTCCGAGGTTTCTGATGACAAT 3809
|||||

```

OY 2758 GGCATCCCATATAATTAATCAAGCCTATTGAGTATGAATACATTCATTCACAGGCTCA 2817
Db 3810 GGCATCCCATATAATTAATCAAGCCTATTGAGTATGAATACATTCATTCACAGGCTCA 3869
OY 2818 GAACACAGCTCTGATCAAGATGATCAACACACAGCTCCGATGGAAACAAACGAGATCTA 2877
Db 3870 GAGCAGAGCTGGATGAGATGATCAAAACAGAGCTCAGAGATTAATAAATCGCATCTA 3929
OY 2878 GTGTATGTTTAATAAATAGGAGATCATTTGTA---GAATTTGCAAGAGGCTCGAAGTG 2933
Db 3930 GTATGTTTACACACATGATTAATAATGATGAGGATTTGTAAGAGGCTCGAATG 3989
OY 2934 CAGGGGTTTTGAAGTTTTGAG---AAAATTATGCAAAATGTCACAGAG-----TTTGTG 2985
Db 3990 CGAGGTGTTTTCAGGTTCTGAGAGTAAATTTATGCAAAATATGTCACAGAGCTATATGTC 4049
OY 2986 GCTCTGTGTACAAATTTATTTCTTCTAGTTATGGAATTTGTTTAAATGTTAAATTT 3045
Db 4050 GCTCTGTGTACAAATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4109
OY 3046 ATTCACCCCTTTTAAATTCAGTAATTTAGAAAAATTTGTTATAGGAAGTAAATATGAA 3105
Db 4110 ATTCACCCGTTTAAATTCAGTAATTTAGAAAAATTTGTTATAGGAAGTAAATATGAA 4169
OY 3106 CTGATATTATAGTCAATTTCTTGGTACTTAAAGTACTTAAAGAGAGAAG 3154
Db 4170 CTGATATTATAGTCAATTTCTTGGTACTTAAAGTACTTAAAGTAAAGTAG 4218

RESULT 5
AF164041 4241 bp mRNA linear PRI 06-DEC-1999
LOCUS Homo sapiens WARTS protein kinase (WARTS) mRNA, complete cds.
DEFINITION AF164041
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4241)
AUTHORS Kishimoto,T., Nawa,S.-i., Nagamine,Y., Nishiyama,Y. and Saya,H.
TITLE WARTS protein, polynucleotide encoding the same, antisense polynucleotide thereof, and antibody recognizing the protein
JOURNAL Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-ku,
Yokohama-shi; Kanagawa;
Japan.;
2. (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiraoka,T., Kitamura,N. and
Saya,H.
TITLE A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
JOURNAL FEBS Lett. 459 (2), 159-165 (1999)
MEDLINE 99447636
PUBMED 10518011
REFERENCE 3 (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto 860-0811,
Japan
FEATURES
source
1. .4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .4241
/gene="WARTS"
33. .3425

```

```

/gene="WARTS"
/Note="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"
/db_xref="GI:5738136"
/translation="MKRSEKPEGYRMPKTPPASNTVTSRQMLQETRESRLNLSRP
SDAAKAEHMSKSTEDPRQVNPFPKFGTHKALQEIINSLPFPANFENSRSTSEVN
POMLODQAAGFDEDMVIALOKTNNRSEIAAIEFISKMSYODPRRQMAAAARPIN
ASMKPGNVQSVNRKQSKGKESLVPQRHGPGLGESVAYHSESPNSDVGRLPES
GISAFVQAHPNSGNRNPVPQVRSVTPPPRGTGTPPPRGTGTPPPRGTGTPPPRGTG
SSKNEVVISISVPVQAQWQEGVPPPLNTSMNPNQSGGSSVVPVGRQPIIMOS
SSKNEVPSGRPMONGTCQDFMILHONVPAQTVNPPPPYPLTAANGQSPSSALQTS
GSAAPSSYNGSIFQSMVMVNRNHNMLYINISVGLQTNWPSQSSAPSSSGHE
ITWOPNIPVRSNFPNPLGNRASHANSQATVTAITPAIQOPVYKSMRVLKPEL
QTALAPHPISWIPQIQTVOPSPFEGTASNTVMPVAEAPNTQPPPPYKRLHLO
NPSVPPYESISKSEKQSLPKDESEKSYENDSGDKKQITSPITVRKKNKDE
ERRSRISQSPQAFKPFMEQHVNLKSHOORLHRKQLENNMRVGLSDQADQMR
KMLCKQESNYIRLAKMDKSMFVKIKITGLIGAFGEVCLARKVDTKALYATKLRKD
VLLRNQVAHVKAERDILAEADNEWVRLYISFQDKNLIYFVMDIIPGGDMSLIRMG
IFPESLARFYIAELTCAVESVHKMGFHRDILKPNILIDRDGHLKLDFTGLCTGRWT
HDSYVQSGDHPQSDMDFSNWGDPSRCGRDLKPLERRAARQHQRLAHSVLGTP
NYIAPEVLLRTGYTLQCDWMSVGLFEMLVGQPPFLAQTLETQVQKVMNNQTSIHIP
POAKLSPEASDLIIKLCRGPEDRLKNGADEIKAHFPFKTIDFSSDLRQQAASVPIKI
THPTDTSNFDVPDVKLWSDNENENVDNLNGYKNGKHPEHAFYETFRPFDDNG
YPYNYPPPIEYIYINSQSEQQSDEDDONTGSEIKNRDLVTV"
/misc_feature
2145. .3059
/gene="WARTS"
/Note="Region: serine/threonine kinase domain"
BASE COUNT 1384 a 884 c 817 g 1156 t
ORIGIN
Query Match 76.9%; Score 2472; DB 9; Length 4241;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 16; Gaps 5;
OY 1 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTCTAGTCTCT 60
Db 534 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTCTAGTCTCT 593
OY 61 CAGACACAGCGGCCATCTCTAGGAGAAAATGTGTTTCTGCTGAAAGCCCAACTCA 120
Db 594 CAGAGGATCGGCCGCCACTAGGAGAAAGTGGCCCTATCTCTGAGAGTCCCACTCA 653
OY 121 CAGCGGATGTAGGAGACCTCTGTCTGATCCGCGATTCGAGCATTTGCTCAAGCTCAC 180
Db 654 CAGACAGATGTAGGAGACCTTTGTCTGATCTGATATATCAGCATTTGTTCAAGCTCAC 713
OY 181 CCAAGCAATGGACAGAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCTCT 240
Db 714 CAGACAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCTCTCTCTCTCT 773
OY 241 CCACCACTCGGAGAGGCGCAGACCCACCTCCCGGAGGAGGAGGAGGAGGAGGAGGAG 300
Db 774 CCACCACTCGGAGAGGCGCAGACCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 833
OY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGACACATGGAGTAGCTATCTCCCGA 360
Db 834 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGACACATGGAGTAGCTATCTCCCGA 893
OY 361 ATCTCCCTGTTCACCTGGGCGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
Db 894 ATCTCTCTCTCCACCTGGGCGTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 953
OY 421 TCTCCATGAATCCCTCTAGCAGGCTCAGAGGGGCAATAGTCTGTCTGTTCCAGTGGTAGA 480
Db 954 TCCCCCATGAATCTCTCTAATCAGGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1013
OY 481 CAACCATCATCATGAGAGTACTAGCAAAATTTAACTTTACACAGGCGGACCTGGAGTT 540
Db 1014 CAACCAATCATCATGAGAGTACTAGCAAAATTTAACTTTTCCATCATGAGGAGACCTGGAATG 1073

```


Qy 541 CAGAAATGGTGGTCACTGATTTTATCGTGCAACAAAAA---TGTCGCCACTGGTTCT 597
 Db 1074 CAGAAATGGTGGTCACTGATTTTATCGTGCAACAAAAA---TGTCGCCACTGGTTCT 1133
 Qy 598 GTGACTGGGAGCCAGCCACTCCATATCTCTGACCCAGCTAATGGCAAGCCCTCT 657
 Db 1134 GTGAATGGGAGCCAGCCACTCCATATCTCTGACCCAGCTAATGGCAAGCCCTCT 1193
 Qy 658 GCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCATTCGCGCAATGGAAAGCTTCTCAG 717
 Db 1194 GCTTTACAAACAGGGGCTTCTGCTGCTCCATATATACAAATGGAAGTATTCCTCAG 1253
 Qy 718 TCGATGATGGTGGCCAGGAGAGCATAAACATGGAGCTTTATAATTAATGTCCT 777
 Db 1254 TCTATGATGGTGGCCAGGAGAGCATAAACATGGAGCTTTATAATTAATGTCCT 1313
 Qy 778 GGACTGCAAAACAGCTGGCCAGCTGCTGCTCTCTGCGCAGTATCCCCAAGCGGT 837
 Db 1314 GGACTGCAAAACAGCTGGCCAGCTGCTGCTCTCTGCGCAGTATCCCCAAGCGGT 1373
 Qy 838 GGGCATGAAATTCCTACATGCGCAACCTAACATACAGTGAAGTCAAAATCTTTTAATAAC 897
 Db 1374 GGGCATGAAATTCCTACATGCGCAACCTAACATACAGTGAAGTCAAAATCTTTTAATAAC 1433
 Qy 898 CCATTAGGAAGTAGAGCAAGTACTCTGCTAAATCTCAGCCCTTCTGCGCAGTACAGTCACT 957
 Db 1434 CCATTAGGAAGTAGAGCAAGTACTCTGCTAAATCTCAGCCCTTCTGCGCAGTACAGTCACT 1493
 Qy 958 GCCATCACCCGCTCCTATTACAGCCCGTGAAGCATGCGCGTCTGCAACCCAGAG 1017
 Db 1494 GCAATACACAGCTCCTATTACAGCCCGTGAAGCATGCGCGTCTGCAACCCAGAG 1553
 Qy 1018 CTGAGAGCTGCTTTAGGCGCCAAACCTCTCTTTGGATGCCACAGCAGTTCAGACTGTT 1077
 Db 1554 CTAGAGCTGCTTTAGGCGCCAAACCTCTCTTTGGATGCCACAGCAGTTCAGACTGTT 1613
 Qy 1078 CAGCTACCCCTTTCTGAGGGTACAGTTCAGTTCAGTGGCTGCTATCCCACTGTTGCT 1137
 Db 1614 CAACCCAGTCTCTTTCTGAGGGAACCGCTTCAAAATGTGACTGTGATGCCACTGTTGCT 1673
 Qy 1138 GAAGCTCAAGCTATCAAGGTCACACCGCTTATCCAAACATCTGCTACACCAAAAC 1197
 Db 1674 GAAGCTCAAGCTATCAAGGTCACACCGCTTATCCAAACATCTGCTACACCAAAAC 1733
 Qy 1198 CCATCTGCTCCCTCATATGAGTCAAGTAAAGCCCTGCAAGATGAACAGCTTAGCTTA 1257
 Db 1734 CCATCTGCTCCCTCATATGAGTCAAGTAAAGCCCTGCAAGATGAACAGCTTAGCTTA 1793
 Qy 1258 CCCAAGGAAGATGATGAGAGAGTGGGACAGTGGTGAATGTTGATAGTGGGATAAAGAAAG 1317
 Db 1794 CCCAAGGAAGATGATGAGAGAGTGGGACAGTGGTGAATGTTGATAGTGGGATAAAGAAAG 1853
 Qy 1318 AAACAGATTAACACTTCACTATCACTGTTGCGAAACAAAGAGATGAAGACGAAGA 1377
 Db 1854 AAACAGATTAACACTTCACTATCACTGTTGCGAAACAAAGAGATGAAGACGAAGA 1913
 Qy 1378 GAGTCTGGATTCAGAGTACTCCCCAGAGCCCTTTAAGTCTTTCATGGAGCAGCAGTA 1437
 Db 1914 GAATCTGATTAACAGTATTCTCTCAAGCATTTAAATCTTTATGGAGCAACATGA 1973
 Qy 1438 GAGAAATCTGAAAGTCTCATCAGCAGCGTCTGATCGGAAGAGAGAGCTAGAAAATGAA 1497
 Db 1974 GAAATCTGAAAGTCTCATCAGCAGCGTCTGATCGGAAGAGAGAGCTAGAAAATGAA 2033
 Qy 1498 ATGATGGGTTGATTTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTGGCAG 1557
 Db 2034 ATGATGGGTTGATTTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTGGCAG 2093
 Qy 1558 AAGAGTCTAACTATATTCGTTTAAAGGCTTAAATGGCAGAGTCTATGTTTGAAG 1617
 Db 2094 AAGAGTCTAACTATATTCGTTTAAAGGCTTAAATGGCAGAGTCTATGTTTGAAG 2153

Qy 1618 ATAAAGACATTAGGAATAGGAGCGTTGGTGAAGTCTGCTAGCAAGAAAGTTCGATACT 1677
 Db 2154 ATAAAGACATTAGGAATAGGAGCGTTGGTGAAGTCTGCTAGCAAGAAAGTTCGATACT 2213
 Qy 1678 AAAGCTTTGATGCAACAAAGACTCTTCGAAGAAAGAGCTTCTGCTCCGAAATCAGGTG 1737
 Db 2214 AAAGCTTTGATGCAACAAAGACTCTTCGAAGAAAGAGCTTCTGCTCCGAAATCAGGTG 2273
 Qy 1738 GCTCATGTGAAGCGGAGAGGATATCTTAGCAAGCGCAATAGTGGTGGTGGTCCGC 1797
 Db 2274 GCTCATGTGAAGCGGAGAGGATATCTTAGCAAGCGCAATAGTGGTGGTGGTCCGC 2333
 Qy 1798 CTGTACTACTCTTCCAGGAGCAAGCAACTTGTACTTTGTGATGGACTACATCTCTGGG 1857
 Db 2334 CTATATTTATTCATTCGAAGTAAAGCAAAATTTATATCTTTGTAATGGACTACATCTCTGGG 2393
 Qy 1858 GGGCATATGATGAGCGCTTAAATAGATGGCATCTTCTGAAATCTCGCAGGATTC 1917
 Db 2394 GGTGATATGATGAGCGCTTAAATAGATGGCATCTTCTGAAATCTCGCAGGATTC 2453
 Qy 1918 TACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTTTATTCATAGA 1977
 Db 2454 TACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTTTATTCATAGA 2513
 Qy 1978 GATATTAACCTGATAACATTTTGTGACCGTGTATGCCATATTAATTAATGACTGACTTT 2037
 Db 2514 GATATTAACCTGATAACATTTTGTGACCGTGTATGCCATATTAATTAATGACTGACTTT 2573
 Qy 2038 GGCCTGTGCACTGGCTTCAGATGGACACATGACTTCCAGTACTACCAGTGGGGATCAC 2097
 Db 2574 GGCCTGTGCACTGGCTTCAGATGGACACATGACTTCCAGTACTACCAGTGGGGATCAC 2633
 Qy 2098 CCAGCGCAAGATACATGGATTTTCAAGTAAAGTGGGAGATCCTTCCAAATGTCGGTGT 2157
 Db 2634 CCAGCGCAAGATACATGGATTTTCAAGTAAAGTGGGAGATCCTTCCAAATGTCGGTGT 2693
 Qy 2158 GGGGACAGACTGAAGCACTGGAGCGGAGAGTCTGCTGCCAGCACCAGCGATGCTAGCC 2217
 Db 2694 GGGGACAGACTGAAGCACTGGAGCGGAGAGTCTGCTGCCAGCACCAGCGATGCTAGCA 2753
 Qy 2218 CATCTCTGTTGGGACTCCCAATTTATTCACCTGAAGTGTACTGCGAAGAGGATAT 2277
 Db 2754 CATCTCTGTTGGGACTCCCAATTTATTCACCTGAAGTGTACTGCGAAGAGGATAT 2813
 Qy 2278 ACACAGTGTGTGACTGGTGGAGTGGTGTATTCTTTTGAAGTGTGGTGGGACAA 2337
 Db 2814 ACACAGTGTGTGACTGGTGGAGTGGTGTATTCTTTTGAAGTGTGGTGGGACAA 2873
 Qy 2338 CCTCTCTTCTTGGCAGAAACCCCATTTAGAACACAAATGAAGTTTATCATCTGGCAACT 2397
 Db 2874 CCTCTCTTCTTGGCAGAAACCCCATTTAGAACACAAATGAAGTTTATCATCTGGCAACT 2933
 Qy 2398 TCTCTACACATCCCTCTCAAGCTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAAA 2457
 Db 2934 TCTCTACACATCCCTCTCAAGCTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAAA 2993
 Qy 2458 CTGTGTCGAGACCAAGAGCCGCTCGCAAGAACCGTGTGATGAGATAAAGGCTCAT 2517
 Db 2994 CTGTGTCGAGACCAAGAGCCGCTCGCAAGAACCGTGTGATGAGATAAAGGCTCAT 3053
 Qy 2518 CCATTTTTTAAAGACATGCTCTCTAGTGTCTGAGACAGAGTCTGCTTATATCATC 2577
 Db 3054 CCATTTTTTAAAGACATGCTCTCTAGTGTCTGAGACAGAGTCTGCTTATATCATC 3113
 Qy 2578 CCTAAATATCAGCATCCCAAGATACATCAATTTCCAGCTCTGTTGATCTGATAAATG 2637
 Db 3114 CCTAAATATCAGCATCCCAAGATACATCAATTTCCAGCTCTGTTGATCTGATAAATG 3173
 Qy 2638 TGGAGCATGGCAGGAGGAGGAAATATCAGTGACACTCTGAGCAGGATGATTAATAAT 2697
 Db 3174 TGGAGCATGGCAGGAGGAGGAAATATCAGTGACACTCTCAATGATGATGATTAATAAT 3233
 Qy 2698 GGGAGCAGCCCGGAGGAGGCTTCTATGATTCACCTTTTCGAGGTTTTTTGATGACAAT 2757

```

Db 3234 GGAAGCATCTGGAACATGCTTATGATTTACCTCCGAAGGTTTTTGTGACAAAT 3293
QY 2758 GGTACCCCATATAATATCAAGCCCTATTGAGTATGAATACATTCATTCACAGGCTCA 2817
Db 3294 GGTACCCCATATAATATCAAGCCCTATTGAGTATGAATACATTCATTCACAGGCTCA 3353
QY 2818 GAACAACAGTCTGATGAAGATGATCAACACACAGCTCCGATGGAACACCCAGAGATCTA 2877
Db 3354 GAGCAGCAGTCCGATCAAGATGATCAACACACAGGCTCAGAGATTAAAAATCGGATCTA 3413
QY 2878 GTGTATGTTTATAAATACTAGGAGATCATTTGTA - - - - GAATTTGCAAGAGGCTGAAGTG 2933
Db 3414 GTATATGTTTAAACACACTAGTAAATAAATGCTATGAGGATTTGTAAGAGGCTGAATG 3473
QY 2934 CAGGGTGTGTTGAAGTTTTGAG - - - AAAATATGCAAAATGTGACAG - - - - TTTGTGT 2985
Db 3474 CGAGGTGTTTGAAGTTTTGAG - - - AAAATATGCAAAATGTGACAG - - - - TTTGTGT 3533
QY 2986 GCTCTGTGACAAATATTTTCTTAAGTTATGGAATTTGTTTAAATGTTAAATTT 3045
Db 3534 GCTCTGTGACAAATATTTTCTTAAGTTATGGAATTTGTTTAAATGTTAAATTT 3593
QY 3046 ATT-CCACCTTTTAAATCAGTAATTTAGAAAAAATTTTAAAGAAAGTAAATATGA 3104
Db 3594 ATTCCAGCGCTTTAAATCAGTAATTTAGAAAAAATTTTAAAGAAAGTAAATATGA 3653
QY 3105 ACTGAGTATATAGTCAATCTTGGTACTTAAAGTACTTAAAGAGAAG 3154
Db 3654 ACTGAATATATAGTCAATCTTGGTACTTAAAGTACTTAAAGAGAAG 3703

```

RESULT 6

E24613 2442 bp DNA linear PAT 18-JUN-2001
 DEFINITION warts protein, polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein.

ACCESSION E24613
 VERSION E24613.1 GI:13024646
 KEYWORDS JP 1999089580-A/1.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 2442)
 AUTHORS Toshihiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
 TITLE warts protein, polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein
 JOURNAL Patent: JP 1999089580-A 1 06-APR-1999;
 COMMENT SUMITOMO ELECTRIC IND LTD
 OS Unidentified
 PN JP 1999089580-A/1
 PD 06-APR-1999
 PF 24-SEP-1997 JP 1997258689
 PR PI TOSHIOHKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
 YASUYUKI NISHIYAMA,
 PI HIDEYUKI SAYA
 PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
 G01N33/532//C12P21/02, (C12P21/02, C12R1:19), C12N15/00,
 PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
 CC C12N15/00, PC C12R1:91
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..2442 /organism='Unidentified'.
 FT Location/Qualifiers
 1..2442 /organism='Unidentified'
 /db_xref='taxon:32644'

FEATURES

source

BASE COUNT

777 a 528 c 491 g 646 t

ORIGIN

```

Query Match 60.6%; Score 1946.8; DB 6; Length 2442;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 292; Indels 4; Gaps 2;
QY 331 TCTGGGAACATGAGTAGTACGTAATCTCCGAATCTCCCTGTTCCACCTGGGGGTGGCAG 390
Db 1 TCTGGGAACATGAGTAGTACGTAATCTCCGAATCTCCCTGTTCCACCTGGGGGTGGCAG 60
QY 391 GAGGGGTACCTCCACACCTCTTACCACCTTCTCCCATGAATCCCTAGCCAGGCTCAG 450
Db 61 GAGGGGTATCTCTCCACACCTTTTCAACACTTCCCTCATGAATCTCTTATCAGACAG 120
QY 451 AGGGCCATTAGTCTCTCTCCAGTTGGTAGAACCCCATCATCATGAGAGTAGTACAGAA 510
Db 121 AGAGGCATTAGTCTCTCTCTCTGTTGGCAGACAAACCAATCATCATGAGAGTTCTAGCAA 180
QY 511 TTAACTTTTACACAGGGGACCTGGAGTTGAGATTGAGTTGGTGGTCTGAGTCTGATTTATC 570
Db 181 TTAACTTTTCCATFCAGGGAGACCTGGAATGCAAGTGGTACTGGACAAATGATTTTCATG 240
QY 571 GTGCACCAAAA - - - TGTCCCACTGTTCTGTGACTCGGCAGCCACCATCTCCATATCCT 627
Db 241 ATACACCAAAATGTTCTGCTCTGCTGCACTGTGAATCGGCAGCCACCATCTCCATATCCT 300
QY 628 CTGACCCAGCTAATGACAAAAGCCCTCTGCTTTTACAAACAGGGCTTCTGCTGCTCCA 687
Db 301 CTGACAGCAGTAAATGACAAAAGCCCTCTGCTTTTACAAACAGGGGATCTGCTGCTCCT 360
QY 688 CCATATTCGCCAANTGGAACGTTCTCTAGTTCGATGATGGTCCCAACAGGAACAGTCAAT 747
Db 361 TCGTCATATACAAATGGAAGTATTTCTCAGTGTATGATGGTCCCAACAGGAATAGTCAAT 420
QY 748 ACATGAGCTTTTAAATATTAATGTCCTGGACTCCAAACAGCCCTGGCCCAAGTCTCT 807
Db 421 ACATGGAATATATTAATGTCCTGGACTCCAAACAGCCCTGGCCCAAGTCTCTCTCTCT 480
QY 808 TCTGCTCTCGCAGTCAATCCCAAGCGTGGGCATGAAATTCCTACATGGCAACCTAAC 867
Db 481 TCTGCTCCAGCCAGTCAATCCCGAGCAGTGGCATGAAATCCCTACATGGCAACCTAAC 540
QY 868 ATACAGTGAAGTCAAAATCTTTTATAACCCATTTAGGAAGTAGAGCAAGTCACTCTGCT 927
Db 541 ATACAGTGAAGTCAAAATCTTTTATAACCCATTTAGGAAGTAGAGCAAGTCACTCTGCT 600
QY 928 AATTCACGCTTCTGCCACTACAGTCACTGCCATCACACCCCTCTTATTCACAGCC 987
Db 601 AATTCACGCTTCTGCTACACAGTCACTGCAATTTACACCCAGTCTCTTATTCACAGCCT 650
QY 988 GTGAAAGCATGCGCTCTGAAACCCAGAGCTGAGACTGCTTTAGCCCAACCCCTCTCT 1047
Db 661 GTGAAAGTATCGCTGCTATTAACCCAGAGCTACAGACTGCTTTAGCACCCTACACACCCT 720
QY 1048 TCTTGGATGCCAGCCAGTTCAGACTGTTACGCTTACCCCTACCCCTTTTCTGAGGGTACAGCT 1107
Db 721 TCTTGGATGCCAGCCAGTTCAGACTGTTTCAACTGTTTCAACCCAGTCTCTTCTGAGGAAACGCT 780
QY 1108 TCAAGTGTGCTCTCATCCACCTGTTGCTGAAGCTCCAAGCTATCAAGTCCACCCAGC 1167
Db 781 TCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAACATATCAAGACCCACCCCA 840
QY 1168 CCTTATCCAAAACATCTGTACACCAAAACCCCTCTGCTCCCTCATATGAGTCAAGTAAAGT 1227
Db 841 CCTTATCCAAAACATCTGTGACCAAAACCCCTCTGCTCCCTCATATGAGTCAAGTAAAGT 900
QY 1228 AGCCCTGCAAGATGAACAGCCCTAGCTTTACCAAGAGATGATAGTGAAGAGTGGG 1287
Db 901 AGCCCTGCAAGATGAACAGCCCTAGCTTTACCAAGAGATGATAGTGAAGAGTGGG 960
QY 1288 GACAGTGGTCACTCTGCGGATTAAGAAAACAGATTAACACTTCACTTCACTCTGTT 1347
Db 961 GAAATGTTGATAGTGGGATTAAGAAAACAGATTAACACTTCACTTCACTCTGTT 1020
QY 1348 CGGAAAAACAAGAGATGAAGAAACGAGAGTCTCGGATTCAGAGTTACTCCCCACAG 1407

```

Db	1021	AGAAAAACAAGAGATGAAGGGAATCTCGTATTCAAAGTATTCTCTCA	1080
Qy	1408	GCCTTTAAAGTCTTTCATGAGCAGCAGCTAGAGAGCGTCTCAAGTCTCATCAGCAGCT	1467
Db	1081	GCATTTAAATCTTTATGAGCAACATGATGAGAAATGTACTCAATCTCATCAGCAGCT	1140
Qy	1468	CTGCATCGAAGACAGCAGCTAGAAATGAATGATCGGGTGGATTAATCTCAAGATGCC	1527
Db	1141	CTACATCGTAAACAAATTTAGAGATGAATGATCGGGTGGATTAATCTCAAGATGCC	1200
Qy	1528	CAGGATCAATGAGAAAGATGCTTTGCCGAGAAAGCTCTAACTATATTCGTCTTAAAGG	1587
Db	1201	CAGATCAATGAGAAAGATGCTTTGCCGAGAAAGATCTAAATACATCCGCTCTTAAAGG	1260
Qy	1588	GCTAAATGGACAAGTCTATGTTGTAAGATGAACACATTAGGATAGAGCGTTTGGT	1647
Db	1261	GCTAAATGGACAAGTCTATGTTGTAAGATGAACACATTAGGATAGAGCGTTTGGT	1320
Qy	1648	GAAGTCTGTCTAGCAAGAAAGTCGATCTAAAGCTTTGATGCAACAAAGACTCTTCGA	1707
Db	1321	GAAGTCTGTCTAGCAAGAAAGTAGATCTAAGCTTTGATGCAACAAAGACTCTTCGA	1380
Qy	1708	AAGAAAGACGCTTCTGCTCCGAAATCAGGTGGCTCATGTGAAACGGAGGAGGATTCCTA	1767
Db	1381	AAGAAAGATGCTTCTTCTCGAAATCAAGTCGCTCATGTTAAGCTGAGAGAGATATCCTG	1440
Qy	1768	GCAGAAGCCGACATGAGTGGTGGTCCGCTGTACTACTCTTCCAGGACAAAGGACAAC	1827
Db	1441	GCTGAAGCTGACATGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1500
Qy	1828	TTGTACTTTGTGATGACTACATCTCTGGGGGGGATATGATGACCTTATTAATAGATG	1887
Db	1501	TTATCTTTGTAAGGACTACATCTCTGGGGGGTATATGATGACCTTATTAATAGATG	1560
Qy	1888	GGCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACTGTGCGAGTTGAA	1947
Db	1561	GGCATCTTTCTGAAAGTCTGGCAGGATCTACATAGCAGAACTTACTGTGCGAGTTGAA	1620
Qy	1948	AGTGTCTCAATAAATGGTGTATTCATAGATATTAACCTGTATAACATTTTATTGATC	2007
Db	1621	AGTGTCTCAATAAATGGTGTATTCATAGATATTAACCTGTATAACATTTTATTGATC	1680
Qy	2008	CCTGATGCCATATTAATGACCTGACCTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2067
Db	1681	CCTGATGCCATATTAATGACCTGACCTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1740
Qy	2068	GACTCCAAAGTACTACAGTGGGATCACCCAGGCAAGATAGCATGATTTTCAATAC	2127
Db	1741	GATTTCTAAGTACTATCAGAGTGGTGACCATCCACGGCAAGATAGCATGATTTTCAATAC	1800
Qy	2128	GAATGGGAGATCTCTTCAATTTGCGGTGGTGGGACAGACTGAAGCCACTGGAGCGGAGA	2187
Db	1801	GAATGGGAGATCTCTTCAATTTGCGGTGGTGGGACAGACTGAAGCCACTGGAGCGGAGA	1860
Qy	2188	GCTGCTCGCCAGCAGCAGGATGCTAGCCCATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGT	2247
Db	1861	GCTGCAAGCCAGCAGCAGGATGCTAGCCCATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGT	1920
Qy	2248	GCACCTGAAGTCTACTCGACAGAGATATACAGCTGTGTGACTGGTGGAGTGTGTGT	2307
Db	1921	GCACCTGAAGTGTGCTACGACAGAGATACACAGTGTGTGATGTTGGTGGTGGTGGTGGTGGT	1980
Qy	2308	GTATTTCTTTTGAATGTTGGTGGGACAACTCTCTTCTTGGCACAACCCCATAGAA	2367
Db	1981	GTATTTCTTTTGAATGTTGGTGGGACAACTCTCTTCTTGGCACAACCCCATAGAA	2040
Qy	2368	ACACAAATGAAGTTATCATCTGGCAAACTTCTTACACATCCCTCCTCAAGTAACTG	2427
Db	2041	ACACAAATGAAGTTATCACTGCAAGTAACTCTTCTTCACTCCACCAAGCTAACTC	2100
Qy	2428	AGTCCTGAAGCTCTGACCTCATTTATCAAACTGTGCGAGGACGAGAACCCGCTCGGC	2487
Db	2101	AGTCTGAAGCTTCTGATCTTATTATAAAGTTCGCCGAGGACCCGAAGATCGGTTAGGC	2160
Qy	2488	AGAAGCGTGTGATGATGAAGCTCATCAATTTTAAAGACCATCAATTTCTCTAGT	2547
Db	2161	AGAAGTGTGTGATGAAGTAAAGCTCATCAATTTTAAACAAATGATTTCTCCAGT	2220
Qy	2548	GATCTGAGACAGCAGCTGCTTCTATACATCCCTAAATACAGCATCCCAACAGATACATCC	2607
Db	2221	GACCTGAGACAGCAGCTGCTTCTATACATCCCTAAATACAGCATCCCAACAGATACATCA	2280
Qy	2608	AATTTTCGACCCCTGTTGATCTCTGATAAATGTTGGAGGATGGAGGAGGAGAAATATC	2667
Db	2281	AATTTTCGACCCCTGTTGATCTCTGATAAATGTTGGAGGATGGAGGAGGAGAAATATCA	2340
Qy	2668	ATGACACTCT-CAGCGGATGTTAAATGGAAGCACCCTGGAGCGCTTCTCTATCA	2726
Db	2341	AATGACACTCTCAATGGGATGTTAAATGGAAGCACCCTGGAGCGCTTCTCTATCA	2400
Qy	2727	GTTCACTTTTCGAGGTTTGTGATGACAAATGGCTACCCATA	2768
Db	2401	ATTACCTTCCGAAGTTTTGTGATGACAAATGGCTACCCATA	2442
RESULT 7			
AL583963/c			
LOCUS	Human DNA sequence from clone RP1-203A15 on chromosome 6, complete	80146 bp	DNA linear PRI 23 JUN-2001
DEFINITION	sequence.		
ACCESSION	AL583963		
VERSION	AL583963.11	GI:14547269	
KEYWORDS	HTG.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 80146)		
JOURNAL	Hammond, S.		
COMMENT	Direct Submission		
	Submitted (23-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,		
	CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk		
	requests: clonerequest@sanger.ac.uk		
	On Jun 25, 2001 this sequence version replaced gi:14133045.		
	During sequence assembly data is compared from overlapping clones.		
	Where differences are found these are annotated as variations		
	together with a note of the overlapping clone name. Note that the		
	variation annotation may not be found in the sequence submission		
	corresponding to the overlapping clone, as we submit sequences with		
	only a small overlap as described above.		
	This sequence was finished as follows unless otherwise noted: all		
	regions were either double-stranded or sequenced with an alternate		
	chemistry or covered by high quality data (i.e., phred quality >=		
	30); an attempt was made to resolve all sequencing problems, such		
	as compressions and repeats; all regions were covered by at least		
	one plasmid subclone or more than one M13 subclone; and the		
	assembly was confirmed by restriction digest. The following		
	abbreviations are used to associate primary accession numbers given		
	in the feature table with their source databases: Em., EMBL; Sw.,		
	SWISSPROT; Tr., TrEMBL; Wp., WORMPEP; information on the WORMPEP		
	database can be found at		
	http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence		
	was generated from part of bacterial clone contigs of human		
	chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping		
	group. Further information can be found at		
	http://www.sanger.ac.uk/HGP/Chr6		
	RP1-203A15 is from the library RPCI-1 constructed by the group of		
	Pieter de Jong. For further details see		
	http://www.chori.org/bacpac/home.htm		
	VECTOR: pCYPAC2		
	IMPORTANT: This sequence is not the entire insert of clone		
	RP1-203A15 It may be shorter because we sequence overlapping		
	sections only once, except for a 100 base overlap.		
	The true left end of clone RP1-317N9 is at 78147 in this sequence.		
	The true right end of clone RP1-703H16 is at 2000 in this		

FEATURES	source	sequence.	Location/Qualifiers
repeat_region			15966..16070
repeat_region			/note="FLAM_C repeat: matches 1..105 of consensus"
repeat_region			16076..16112
repeat_region			/note="Alu repeat: matches 79..160 of consensus"
repeat_region			16122..17029
repeat_region			/note="L1MA4A repeat: matches 5319..6300 of consensus"
repeat_region			17080..17261
repeat_region			/note="MER74A repeat: matches 1..171 of consensus"
repeat_region			17490..17788
repeat_region			/note="Alusg repeat: matches 1..296 of consensus"
repeat_region			17966..18263
repeat_region			/note="Aluy repeat: matches 1..298 of consensus"
repeat_region			18502..18551
repeat_region			/note="25 copies 2 mer ta 100% conserved"
repeat_region			18552..18818
repeat_region			/note="Alusp repeat: matches 34..303 of consensus"
repeat_region			18929..19117
repeat_region			/note="MIR repeat: matches 41..241 of consensus"
repeat_region			19252..19557
repeat_region			/note="Alusg repeat: matches 1..306 of consensus"
repeat_region			19558..19843
repeat_region			/note="AluJo repeat: matches 12..294 of consensus"
repeat_region			20971..21270
repeat_region			/note="Alusx repeat: matches 1..300 of consensus"
repeat_region			21302..21391
repeat_region			/note="FRAM/FAM repeat: matches 4..72 of consensus"
repeat_region			22015..22173
repeat_region			/note="MER5B repeat: matches 7..176 of consensus"
repeat_region			22419..23403
repeat_region			/note="Cpg island"
repeat_region			22531..22618
repeat_region			/note="44 copies 2 mer cc 64% conserved"
repeat_region			23378..23453
repeat_region			/note="MER5A repeat: matches 108..185 of consensus"
repeat_region			23686..23896
repeat_region			/note="Alusg/x repeat: matches 89..310 of consensus"
repeat_region			24365..24495
repeat_region			/note="FLAM_C repeat: matches 2..125 of consensus"
repeat_region			24659..24859
repeat_region			/note="MER3 repeat: matches 2..209 of consensus"
repeat_region			25072..25249
repeat_region			/note="Alusx repeat: matches 118..290 of consensus"
repeat_region			25256..25572
repeat_region			/note="Alusx repeat: matches 1..309 of consensus"
repeat_region			25575..25874
repeat_region			/note="Alusx repeat: matches 1..299 of consensus"
repeat_region			26374..26417
repeat_region			/note="22 copies 2 mer ta 100% conserved"
repeat_region			27253..27549
repeat_region			/note="Alusg repeat: matches 1..298 of consensus"
repeat_region			27551..27631
repeat_region			/note="LIME repeat: matches 5699..5778 of consensus"
repeat_region			27852..27915
repeat_region			/note="L2 repeat: matches 2203..2267 of consensus"
repeat_region			28020..28320
repeat_region			/note="Alusx repeat: matches 1..305 of consensus"
repeat_region			28891..29184
repeat_region			/note="Alusx repeat: matches 1..292 of consensus"
repeat_region			29260..29548
repeat_region			/note="Alusg repeat: matches 1..305 of consensus"
repeat_region			29567..29652
repeat_region			/note="AluJo/FAM repeat: matches 202..287 of consensus"
repeat_region			29818..30056
repeat_region			/note="Alusg repeat: matches 1..293 of consensus"
repeat_region			30267..30570
repeat_region			/note="Alusx repeat: matches 1..302 of consensus"
repeat_region			30581..30762
repeat_region			/note="LIMEC repeat: matches 2353..2196 of consensus"
repeat_region			30827..31119
repeat_region			/note="AluJo repeat: matches 1..312 of consensus"
repeat_region			31127..31259
repeat_region			/note="LIMEC repeat: matches 2112..2246 of consensus"
repeat_region			1298..1421
repeat_region			/note="L2 repeat: matches 2234..2359 of consensus"
repeat_region			2135..2886
repeat_region			/note="Cpg island"
repeat_region			2728..3050
repeat_region			/note="AluJo repeat: matches 1..303 of consensus"
repeat_region			3174..3476
repeat_region			/note="Aluy repeat: matches 2..303 of consensus"
repeat_region			3512..3956
repeat_region			/note="L2 repeat: matches 2264..2750 of consensus"
repeat_region			3975..4127
repeat_region			/note="Alusg/x repeat: matches 145..297 of consensus"
repeat_region			4132..4469
repeat_region			/note="Alusx repeat: matches 1..304 of consensus"
repeat_region			4470..4856
repeat_region			/note="HUPRS-P3 repeat: matches 6797..7185 of consensus"
repeat_region			4850..6348
repeat_region			/note="HUPRS-P3 repeat: matches 5120..5625 of consensus"
repeat_region			6349..6647
repeat_region			/note="Alusg repeat: matches 1..301 of consensus"
repeat_region			6648..6815
repeat_region			/note="HUPRS-P3 repeat: matches 4954..5120 of consensus"
repeat_region			6840..6911
repeat_region			/note="HUPRS-P3 repeat: matches 4691..4762 of consensus"
repeat_region			6912..6953
repeat_region			/note="LIPBa repeat: matches 742..701 of consensus"
repeat_region			6974..7347
repeat_region			/note="HUPRS-P3 repeat: matches 4568..4942 of consensus"
repeat_region			7349..7384
repeat_region			/note="LIP repeat: matches 1790..1825 of consensus"
repeat_region			7382..7745
repeat_region			/note="LIPBa repeat: matches 1590..1958 of consensus"
repeat_region			7745..8418
repeat_region			/note="LIM2 repeat: matches 84..638 of consensus"
repeat_region			8458..8921
repeat_region			/note="LIM2 repeat: matches 677..198 of consensus"
repeat_region			9470..9751
repeat_region			/note="Alusx repeat: matches 12..293 of consensus"
repeat_region			9941..10043
repeat_region			/note="AluJo repeat: matches 210..306 of consensus"
repeat_region			10059..10260
repeat_region			/note="Alusg/x repeat: matches 108..309 of consensus"
repeat_region			10295..10994
repeat_region			/note="LIPB1 repeat: matches 4004..4703 of consensus"
repeat_region			10995..11289
repeat_region			/note="Alusp repeat: matches 1..301 of consensus"
repeat_region			11290..12343
repeat_region			/note="LIPB1 repeat: matches 4703..5780 of consensus"
repeat_region			12344..12479
repeat_region			/note="Alusx repeat: matches 1..142 of consensus"
repeat_region			12480..12769
repeat_region			/note="Aluy repeat: matches 1..290 of consensus"
repeat_region			12770..12957
repeat_region			/note="Alusx repeat: matches 142..309 of consensus"
repeat_region			12958..13331
repeat_region			/note="LIPB1 repeat: matches 5779..6152 of consensus"
repeat_region			13382..13962
repeat_region			/note="L1MA4A repeat: matches 3027..3603 of consensus"
repeat_region			13963..14255
repeat_region			/note="Alusc repeat: matches 1..306 of consensus"
repeat_region			14256..15963
repeat_region			/note="L1MA4A repeat: matches 3603..5331 of consensus"

[illegible]

RESULT 8	
LOCUS	E24614
DEFINITION	1374 bp DNA linear PAT 18-JUN-2001
	warts protein, polynucleotide encoding the protein, antisense
	polynucleotide thereof and antibody recognizing the protein.
ACCESSION	E24614
VERSION	E24614.1 GI:13024647
KEYWORDS	JP 199089580-A/2.
SOURCE	unidentified.
ORGANISM	unidentified.
REFERENCE	1 (bases 1 to 1374)
AUTHORS	Toshihiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
TITLE	warts protein, polynucleotide encoding the protein, antisense
	polynucleotide thereof and antibody recognizing the protein
JOURNAL	Patent: JP 199089580-A 2 06-APR-1999;
	SUMITOMO ELECTRIC IND LTD
COMMENT	OS Unidentified
	PN JP 199089580-A/2
	PD 06-APR-1999
	PF 24-SEP-1997 JP 1997258689
	PR
	PI TOSHIIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
	YASUYUKI NISHIYAMA,
	PI HIDEYUKI SAYA
	PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
	G01N33/532//C12P21/02
	PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
	(C12N15/00, PC C12R1:91)
	CC Strandedness: Double;

CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1. 1374
FT /organism="Unidentified"
FEATURES
Location/Qualifiers
1. 1374
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 442 a 252 c 288 g 392 t
ORIGIN

Query Match 35.1%; Score 1126.8; DB 6; Length 1374;
Best Local Similarity 89.2%; Pred. No. 2.6e-261;
Matches 1226; Conservative 0; Mismatches 147; Indels 1; Gaps 1;

Y 1396 TACTCCACAGCGCTTAAAGTCTTCAATGAGCAGCAGCTAGAGAGCTCCTGAAGTCT 1455
D 1 TATCTCTCAAGCATTTAAATCTTTATGAGCAACATGTAGAAATGTACTCAATCT 60

Y 1456 CATCAGAGCGTCTGCATCGGAAGAAGCAGCTAGAAATGAATGATCGGGTGGATTA 1515
D 61 CATCAGAGCGTCTACATCGTAAAAACAATTAGAGATGAATGATCGGGTGGATTA 120

Y 1516 TCTCAAGTCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAATATATT 1575
D 121 TCTCAAGTCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGATCTAATATATT 180

Y 1576 CGTCTTAAAGGCTTAAATGGACAGTCTATGTTTGTAAAGATAAAGACATTAGGATA 1635
D 181 CGTCTTAAAGGCTTAAATGGACAGTCTATGTTTGTAAAGATAAAGACATTAGGATA 240

Y 1636 GGAGCGTTTGTGAGTCTGTAGCAAGAAAGTCTGATCTAAGCTTTGTATGCAACA 1695
D 241 GGAGCATTTGTGAGTCTGTAGCAAGAAAGTCTGATCTAAGCTTTGTATGCAACA 300

Y 1696 AAGACTCTTGAAGAAAGAGCGTTCTGCTCCGAAATAGGTGGTCTATGTGAAAGCGGAG 1755
D 301 AAAACTCTTGAAGAAAGAGTCTGCTCCGAAATAGGTGGTCTATGTGAAAGCGGAG 360

Y 1756 AGGATATCTTGAAGAGCGGAGCAATGAGTGGTGGTCCGCTTACTACTCTTCCAG 1815
D 361 AGGATATCTTGAAGAGCGGAGCAATGAGTGGTGGTCCGCTTACTACTCTTCCAG 420

Y 1816 GACAAGGACAACTTGTACTTGTGATGAGTACATCTTCTGGGGGGATATGATGAGCCTA 1875
D 421 GATAAGGACAACTTGTACTTGTGATGAGTACATCTTCTGGGGGGATATGATGAGCCTA 480

Y 1876 TTAATTAGATGGCATCTTCTTGAATCTGGCAGATTTCTACATAGCAACTTACC 1935
D 481 TTAATTAGATGGCATCTTCTTGAATCTGGCAGATTTCTACATAGCAACTTACC 540

Y 1936 TGTGAGTTGAAGTGTTCATAAATGGTGTATTCATAGAGATATTAACCTGTAAAC 1995
D 541 TGTGAGTTGAAGTGTTCATAAATGGTGTATTCATAGAGATATTAACCTGTAAAC 600

Y 1996 ATTTGATTTGACCTGATGGCCATTAATAATGACTGCTTGGCTTGTGCACTGGCTTC 2055
D 601 ATTTGATTTGACCTGATGGCCATTAATAATGACTGCTTGGCTTGTGCACTGGCTTC 660

Y 2056 AGATGGACACATGACTCAAGTACTACAGATGGGATACCCACCGCAAGATAGCATG 2115
D 661 AGATGGACACATGACTCAAGTACTACAGATGGGATACCCACCGCAAGATAGCATG 720

Y 2116 GATTTCACTAAGTGGGAGATCTTCCAAATGCTGGTGGGAGACAGACTCAAGCCA 2175
D 721 GATTTCACTAAGTGGGAGATCTTCCAAATGCTGGTGGGAGACAGACTCAAGCCA 780

Y 2176 CTGAGCGGGAGAGTCTGCGCAGCAGCAGGATGCTTAGCCCATCTCTGTTGGGACT 2235
D 781 TTAGCGGGAGAGTCTGCGCAGCAGCAGGATGCTTAGCCCATCTCTGTTGGGACT 840

Y 2236 CCCAATTATATGCACTGAAAGTCTACTGCGAAGACAGATATACAGCTGTGTGACTGG 2295
D 2295

Db 841 CCCAATTATATGACCTGAAGTGTGCTACGAACAGGATACACACAGTCTGTGATTGG 900
Qy 2296 TGGAGTGTGTGTTATTTCTTTTGAATGTTGGTGGACAACTCTCTTTCTTGGCACA 2355
Db 901 TGGAGTGTGTGTTATTTCTTTTGAATGTTGGTGGACAACTCTCTTTCTTGGCACA 960
Qy 2356 ACCCATTTAGAAACACAAATGAAGTGTATCATCTGGCAAACTTCTCTACATCTCTCT 2415
Db 961 ACACCATTTAGAAACACAAATGAAGTGTATCACTGGCAAACTTCTCTACATCTCTCA 1020
Qy 2416 CAAGTAAGTGTGCTTGAAGCTCTGACCTCATATTAACAATGTTGTCGAGGACAGAA 2475
Db 1021 CAAGTAACTCAGTCTCTGAAAGTCTGATCTTATTTAACTTTGCGGAGGACCCGAA 1080
Qy 2476 GACCGCTCGCAAGAGGCTGATGAGATAAGGCTCATCATTTTAAAGACCATC 2535
Db 1081 GATCGCTTAGGCAAGATGGTGTGATGAAATAAAGCTCATCTTTTAAACAAAT 1140
Qy 2536 GATTTCTAGTGTCTGAGACAGCAGTCTCTTCTCATACATCCCTAAATCAGGATCA 2595
Db 1141 GATTTCTAGTGTCTGAGACAGCAGTCTCTTCTCATACATCCCTAAATCAGGATCA 1200
Qy 2596 ACAGATACATCAATTTGACCCCTGTTGATCTCTGATAAATTTGTCGAGGAGTGG 2655
Db 1201 ACAGATACATCAATTTGATCTCTGTTGATCTCTGATAAATTTGTCGAGGAGTGG 1260
Qy 2656 GAGAAATATACAGTACATCTCTGAGGAGTGTATTAATAATGGGAGGACCCCGAGCA 2714
Db 1261 GAGAAATATACAGTACATCTCTGAGGAGTGTATTAATAATGGGAGGACCCCGAGCA 1320
Qy 2715 CGCTTCTATGATGATCTTCTGAGGCTTTTGTGATGACAAATGGCTACCCATA 2768
Db 1321 TGCATTTCTAGTGTCTTCTGAGGCTTTTGTGATGACAAATGGCTACCCATA 1374

RESULT 9
AC102712/c 164490 bp DNA linear HTG 21-AUG-2002
LOCUS Mus musculus clone RP24-258P4, WORKING DRAFT SEQUENCE, 17 unordered
DEFINITION pieces.
AC102712
VERSION AC102712.2 GI:22381704
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 164490)
AUTHORS Birren, B., Nusbaum, C. and Lander, E.
TITLE Mus musculus, clone RP24-258P4
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 164490)
AUTHORS Anderson, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choquel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Farreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gordon, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamaze, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Olliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164490)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguski,Y., Bouckgaert,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Macdonald,P., Major,J., Matthews,C.,
Liu,G., Maclean,C., Macdonald,P., Mayhew,T., Mlenka,V.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenka,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
PhunKhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testfay,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17060822.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L19333
Center clone name: 258.P.4

Sequencing Summary Statistics
Chemistry: Dye-terminator Big Dye; n/a; 100% of reads
Assembly: Phrap; version 0.960731
Consensus quality: 158929 bases at least Q40
Consensus quality: 161241 bases at least Q30
Consensus quality: 162174 bases at least Q20
Insert size: 176000; agarose-fp
Quality coverage: 7.2 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 17 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1303: contig of 1303 bp in length
1304 1403: gap of 100 bp
1404 2163: contig of 760 bp in length
2164 2263: gap of 100 bp
2264 3531: contig of 1268 bp in length
3532 3631: gap of 100 bp
3632 4716: contig of 1085 bp in length
4717 4816: gap of 100 bp
4817 8090: contig of 3274 bp in length
8091 8190: gap of 100 bp
8191 11116: contig of 2926 bp in length
11117 11216: gap of 100 bp
11217 41036: contig of 29820 bp in length
41037 41136: gap of 100 bp
41137 45062: contig of 3926 bp in length

* 45063 45162: gap of 100 bp
45163 48629: contig of 3467 bp in length
48630 48729: gap of 100 bp
48730 57775: contig of 9046 bp in length
57776 57875: gap of 100 bp
57876 70851: contig of 12976 bp in length
70852 70951: gap of 100 bp
70952 85997: contig of 15046 bp in length
85998 86097: gap of 100 bp
86098 104897: contig of 18800 bp in length
104898 104997: gap of 100 bp
104998 120197: contig of 15200 bp in length
120198 120297: gap of 100 bp
120298 136805: contig of 16508 bp in length
136806 136905: gap of 100 bp
136906 160104: contig of 23199 bp in length
160105 160204: gap of 100 bp
160205 164490: contig of 4286 bp in length.

Location/Qualifiers
1. 164490
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RP24-258P4"
/clone="RP24-258P4"

1. 1303
/note="assembly_fragment"
1404. 2163
/note="assembly_fragment"
2264. 3531
/note="assembly_fragment"
3632. 4716
/note="assembly_fragment"
4817. 8090
/note="assembly_fragment"
8191. 11116
/note="assembly_fragment"
11217. 41036
/note="assembly_fragment"
41137. 45062
/note="assembly_fragment"
45163. 48629
/note="assembly_fragment"
48730. 57775
/note="assembly_fragment"
57876. 70851
/note="assembly_fragment"
70952. 85997
/note="assembly_fragment"
86098. 104897
/note="assembly_fragment"
104998. 120197
/note="assembly_fragment"
120298. 136805
/note="assembly_fragment"
136906. 160104
/note="assembly_fragment"
160205. 164490
/note="assembly_fragment"
clone_end:77
vector_side:right

BASE COUNT 47330 a 35286 c 34742 g 45505 t 1627 others
ORIGIN

Query Match 33.6%; Score 1079.4; DB 2; Length 164490;
Best Local Similarity 99.8%; Pred. No. 9.6e-250;
Matches 1091; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GTGCAACATTCAATTAAACCAAAAGCTGGAAGCTTCTAAAGAGTCTTAGTTCTCT 60
|||||
Db 5913 GTGCAACATTCAATTAAACCAAAAGCTGGAAGCTTCTAAAGAGTCTTAGTTCTCT 5854
|||||
QY 61 CAGACACGCGCCCTCTCTAGGAGAAATGTGGTTTATCGTTCTTGAAGCCCCCACTCA 120
|||||

Db 5853 CAGAGACACGGCCCATCTCTAGGAGAAAATGTGTTTATCTGTCTCTGAAAGCCCAACTCA 5794
 Oy 121 CAGCGGATGTAGGAAGACCTCTGTCTGATCCGCGCATTTGACAGATTTGCTCAAGCTCAC 180
 Db 5793 CAGCGGATGTAGGAAGACCTCTGTCTGATCCGCGCATTTGACAGATTTGCTCAAGCTCAC 5734
 Oy 181 CAAGCAATGTAGGAGAGTGAACCCCGCCACACACCTCAAGTTAGGAGTGTACTCCT 240
 Db 5733 CAAGCAATGTAGGAGAGTGAACCCCGCCACACACCTCAAGTTAGGAGTGTACTCCT 5674
 Oy 241 CCACCACTCCGAGAGGCGACACCCACCTCCCGAGGACACCTCCCGCTCCCGCTCA 300
 Db 5673 CCACCACTCCGAGAGGCGACACCCACCTCCCGAGGACACCTCCCGCTCCCGCTCA 5614
 Oy 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAACATGAGTACGTAATCTCCCGA 360
 Db 5613 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAACATGAGTACGTAATCTCCCGA 5554
 Oy 361 ATCTCCCTGTTCACCTTGGGCGTGGCAGGAGGGTACCCCTCCACCTCTTACCCT 420
 Db 5553 ATCTCCCTGTTCACCTTGGGCGTGGCAGGAGGGTACCCCTCCACCTCTTACCCT 5494
 Oy 421 TCTCCCATGAATCCCGCTAGCAGGCTCAGAGGCGCATAGTCTGTTCAGTGTGTAGA 480
 Db 5493 TCTCCCATGAATCCCGCTAGCAGGCTCAGAGGCGCATAGTCTGTTCAGTGTGTAGA 5434
 Oy 481 CAACCACTCATCAGAGTACTAGCAAAATTTAACTTTACACAGGCGGACCTCGAGTT 540
 Db 5433 CAACCACTCATCAGAGTACTAGCAAAATTTAACTTTACACAGGCGGACCTCGAGTT 5374
 Oy 541 CAGAAATGTGTGTGCTGATTTATTCGTGTCACCAAAATGTCCTCCACTGTTCTGTG 600
 Db 5373 CAGAAATGTGTGTGCTGATTTATTCGTGTCACCAAAATGTCCTCCACTGTTCTGTG 5314
 Oy 601 ACTCGGACGACACCTCCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660
 Db 5313 ACTCGGACGACACCTCCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 5254
 Oy 661 TTACAAACAGGGGCTTCTGCTCTGCTCAGCATATTCGCAATGGAAGCTTCTCAGTCG 720
 Db 5253 TTACAAACAGGGGCTTCTGCTCTGCTCAGCATATTCGCAATGGAAGCTTCTCAGTCG 5194
 Oy 721 ATGATGTGTGTCACAGGACAGTCAATACATGAGCTTTATATTAATGTCCTCGGA 780
 Db 5193 ATGATGTGTGTCACAGGACAGTCAATACATGAGCTTTATATTAATGTCCTCGGA 5134
 Oy 781 CTGCAACAGGCTGCGCCAGTCTGCTCTGCTGCGCAGTCAATCCCAAGCGGTGG 840
 Db 5133 CTGCAACAGGCTGCGCCAGTCTGCTCTGCTGCGCAGTCAATCCCAAGCGGTGG 5074
 Oy 841 CATGAAATTCCTACATGCAACCTAACATACAGTGAAGTCAATTCCTTTAATACCCA 900
 Db 5073 CATGAAATTCCTACATGCAACCTAACATACAGTGAAGTCAATTCCTTTAATACCCA 5014
 Oy 901 TTAGGAAGTAGGCAAGTCACTCTGTAATTCAGCCCTTCTGCACTACAGTCACTGCT 960
 Db 5013 TTAGGAAGTAGGCAAGTCACTCTGTAATTCAGCCCTTCTGCACTACAGTCACTGCT 4954
 Oy 961 ATCAGACCGCTCTCTATTCAACAGCGCGTGAACATGCGGCTCTGAAACCAAGCTG 1020
 Db 4953 ATCAGACCGCTCTCTATTCAACAGCGCGTGAACATGCGGCTCTGAAACCAAGCTG 4894
 Oy 1021 CAGACTGCTTTAGCCCCCAACCATCTCTTCTGATGCGACAGCCAGTTCAGACTGTCAG 1080
 Db 4893 CAGACTGCTTTAGCCCCCAACCATCTCTTCTGATGCGACAGCCAGTTCAGACTGTCAG 4835
 Oy 1081 CCTACCCCTTTT 1093
 Db 4834 CCTACCCCTTTT 4822

RESULT 10
 AC102712

LOCUS AC102712 164490 bp DNA linear HTG 21-AUG-2002
 DEFINITION Mus musculus clone RP24-258P4, WORKING DRAFT SEQUENCE, 17 unordered pieces.
 ACCESSION AC102712
 VERSION AC102712.2 GI:22381704
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Nusbaum,C. and Lander,E.
 1 (bases 1 to 164490)
 TITLE Mus musculus, clone RP24-258P4
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 164490)
 AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouknight,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginder,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarty,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Riese,C., Rogov,P., Roman,J., Retta,R., Rieback,M., Riley,R., Riese,C., Rogov,P., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausman,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viela,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164490)
 AUTHORS Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Bouknight,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Riese,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viela,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060822.
 All repeats were identified using RepeatMasker:
 Smit,A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L19333

[illegible]

QY 1283 GTGCGACAGTGGTACTCTGGGGATAAGAAAAAGAGAGATTAACAATTCACCTATCA 1342
 Db 137629 GTGCGACAGTGGTACTCTGGGGATAAGAAAAAGAGAGATTAACAATTCACCTATCA 137688
 QY 1343 CTGTTCCGAAACAAAGAAAGATGAAGAACAGAGAGTCTCGGATTCAGAGTTACTGCC 1402
 Db 137689 CTGTTCCGAAACAAAGAAAGATGAAGAACAGAGAGTCTCGGATTCAGAGTTACTGCC 137748
 QY 1403 CACAGGCTTTAAAGTCTTCATGGAGCAGCAGTACAGAACGCTCTGAAGTCTCATCAGC 1462
 Db 137749 CACAGGCTTTAAAGTCTTCATGGAGCAGCAGTACAGAACGCTCTGAAGTCTCATCAGC 137808
 QY 1463 ACGCTCTCATCGGAAGAGCAGCTAGAAAAATGAATGATGCGGGT 1508
 Db 137809 ACGCTCTCATCGGAAGAGCAGCTAGAAAAATGAATGATGCGGGT 137854

RESULT 11
 AR201457 3155 bp DNA linear PAT 20-APR-2002
 LOCUS AR201457
 DEFINITION Sequence 7 from patent US 6359193.
 ACCESSION AR201457
 VERSION AR201457.1 GI:20252345
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3155)
 AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
 TITLE Nucleotide sequences of lats genes.
 JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
 FEATURES Location/Qualifiers
 source 1..3155 /organism="unknown"

BASE COUNT 751 a 924 c 894 g 586 t
 ORIGIN
 Query Match 25.3%; Score 812.6; DB 6; Length 3155;
 Best Local Similarity 65.1%; Pred. No. 2.5e-185;
 Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 919 CACTCTGTCTAATTCACAGCTTCTGCCACTACAGTCACTGCCATCACACCCGCTCTATT 978
 Db 985 CAACCTGAGCCCTACTGCCGCCCCCAACAGCGTCAACCGCTGACGCCGCCACACATC 1044
 QY 979 CAACAGCCGTTGAAGAGATGGCGTCTCTGAACAGAGAGTCAAGTCTGTTTACCCCA 1038
 Db 1045 CTTACCTCTGTAAGAGCGTGGTGTGTGGGCCGAGCCGCCAGACAGCGCGTGGGCC 1104
 QY 1039 ACCCATCTCTTGGATGCCAGCCAGTTCAGACTGTTACGCTACCCCTACCCCTTTTCTGAG 1098
 Db 1105 TGCACCCCGCTGGTGGCTGCGGCCCAACAGCACCTGCCACTGAGAGGCTGGAGACGAG. 1164
 QY 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTGTGCTGAAGCTTCAAGCTATCAAGGT. 1158
 Db 1165 GAGGCGAGCGAGGCCACACCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGTGC 1224
 QY 1159 CCACACCCGCTTATCCAAACATCTGTACACCAAAACCCATCTGTCCTCCATATAG 1218
 Db 1225 CCACCGCTCCGCTATCCAAAGCATTGCTGTGCTGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
 QY 1219 TCAGTAAGTAAAGCTGCAAGATCAACAGCTACCTTACCCCAAGAGATCATAGTAG 1278
 Db 1285 GACCTGGACAGCTGTGCACCAAGTGTGCAGAGAGTCTGCGAGGGGCGACTGATCAGAC 1344
 QY 1279 AAGAGTGGCGGACAGT-----GGTGACTCTGGGGATAAAGAAAGAAACAG 1323
 Db 1345 GCGAGTGACAGAGCCCAAGAGTGGAGGAGGAGACAAAGCTGGCAGAGACAAAGACAG 1404
 QY 1324 ATTACAACTTACCTATCAGTCTTTCGGAACAAAGAGATGAAGAGACGAGAGTCT 1383
 Db 1405 ATTACAGCTTCCCGGCTGCTGTCCGCAAGATAGCAGATGAAGAGAGAGAGTCT 1464

QY 1384 CGGATTCAGAGTTACTTCCCCACAGCCCTTAAAGTTTCTTATGAGCAGCAGCAGTACAGAAC 1443
 Db 1465 CGCATCAGAGTTACTTCCCTTATCCCTTCAATTTCTTATGAGCAGCAGCAGTGGAGAT 1524
 QY 1444 GTCTTGAAGTCTCATCAGCAGCGTCTGATCGGAAGAGAGCAGCTAGAAAATGAATGATG 1503
 Db 1525 GTCATCAAAACCTACACAGCAGAGGTTCAGCGGAGGTTCAGCTGAGCAGGAAATGGCC 1584
 QY 1504 CGGTTGGATTATCATCAGATGCCAGGATCAATTCAGAAAGATGCTTTGCCAGAAAGAG 1563
 Db 1585 AAAGCTGGGCTCTGTGAGGCCCGCAGCAGGAGATGAGGAAGATCTCTACCAAGAGAG 1644
 QY 1564 TCTAACTATATTCGCTTTAAAGGGCTAAATGGACAAAGTCTATCTTTGTAAAGATAAG. 1623
 Db 1645 TCTAACTACACCGGCTGAAGAGGGCCAGATGGACAAAGTCCATGTTGTGAAATCAAG 1704
 QY 1624 ACATTAGGAATAGGAGCGTTTGGTGAAGTCTCTCTAGCAAGAAAGTGCATCTACTAAAGCT. 1683
 Db 1705 ACTCTAGCATCGGTGCGCTTTGGGGAAGTGTGCTTCTGCTCCGAAATCAGGTGCTCAT 1764
 QY 1684 TTGATGCAACAAAGACATCTTCGAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGCTCAT 1743
 Db 1765 CTGTAGGCCATGAGACTCTCAGGAAGAGAGGATGCTCTGAACCGGAATCAAGTGGCCAT 1824
 QY 1744 GTGAAGCGGAGAGGATATCTCTAGCAAGCGCAGCAATGAGTGGTGGTCCGCTGTAC 1803
 Db 1825 GTCAAGGCTGAGAGGAGATCTCTGCTGAAGCAGCAATGAGTGGTGGTCAAACTTAC 1884
 QY 1804 TACTCTTTCCAGGCAAGCAAGCAAACTTGTACTTTGTGATGGACTACATCTCTGGGGGAT 1863
 Db 1885 TACTCTTTCCAGGCAAGCAAGCAAGCTGTACTTTGTGATGGACTACATACCAGCGGAT 1944
 QY 1864 ATGATGAGCCTTAAATAGAAATGGGATCTTCTCTGAAATCTGGCAGATTTACATA 1923
 Db 1945 ATGATGAGCCTGCTGATCAGGATGGAGTCTTCCCTGAGCAGCTGCGCCGCTTCTACAT 2004
 QY 1924 GCAGAACTTACTCTGAGTTCAGAAAGTGTTCATAAAATGGGTTTATTCATACAGATAT 1983
 Db 2005 GCAGAGTTCAGCTGGCCATTTGAAAGTGTCCAAAGTGGGCTTTATCCCGGAGACATC 2064
 QY 1984 AAACCTGATAACATTTTGTATGACCGTGTATGGCCATATTAATTCAGTCTGTTGGCTG 2043
 Db 2065 AAGCTTGACACATATCTATCGACCTGATGTCATATTAAGTGCAGATTTTGGCCTC 2124
 QY 2044 TGCACTGGCTTCAGATGACACATGACTTCAAGTACTACAGAGTGGGATACCCAGCG 2103
 Db 2125 TGCACTGGATTCAGTGGACTCACAAATTCAGTACTACCAAGAGGAAAGGACACATGAGA 2184
 QY 2104 CAAGATAGCATGGATTTCAGTAACGAATGGGAGATCTTCCAAATTTGCGTGTGGGAC 2163
 Db 2185 CAGGACAGTATGGAGCCCGGTGACCTCTGGGACGATGTTTCACTGCTGTGGAGAC 2244
 QY 2164 AGACTGAAGCCACTGGAGCGGAGAGTCTGCTGCCAGCAGCAGCAGTGTGTAGCCCATTT 2223
 Db 2245 AGTTAAAGCCCTGGAGCAGAGGCGCAGAGCAGCAGAGAGTGTGCTGCACATTTCT 2304
 QY 2224 CTGGTTGGACTTCCCAATATTTATTCACCTGAAAGTGTACTGCGAAGAGGATATACAG 2283
 Db 2305 CTGTGCGGACACCAATATACATCTCGGAGGTGCTTCTCCGCAAGAGGTTACACGAG 2364
 QY 2284 CTGTGTACTGTGGAGTGTGGTGTATTTCTTGTGAATGTTGTTGGGAGCAACCTCT 2343
 Db 2365 CTCTGTACTGTGGAGCGTGGTGTGATTTCTTTTGTAGATGCTGGTGGGAGCGGCT 2424
 QY 2344 TTCTTGGCACAACCCCAATTAAGAACACAAATGAAGGTTATCATCTGCAAACTTCTCTA 2403
 Db 2425 TTCTTGGCCCCCACCACAGAGCAGCTGAAGGTGATCACTGGGAGGACGCTG 2484
 QY 2404 CACATCTCTCTCAAGCTTAAGTGTAGTCTGAAGCCTCTGACCTCATTTATCAAACTGTGT 2463
 Db 2485 CATATCTCTAGCAGGTGAGGCTCAGCGCTGAGGCGCCGAGACCTCATCACGAAGCTGTGC 2544

QY 2464 CGAGGACCAAGACCCCTCGCGCAGACAGCGGTGCTGATGAGATAAAGGCTCATCCATT 2523
 Db 2545 TGCCCGGCTGACTGCGCCCTGGCGGAGGATGGCGAGATGACCTCAAGGCACACCCGTTT 2604
 QY 2524 TTTAAGACCATCATTTCTCTAGTATCTGAGACAGCTCTCTCATACATCCCTTAA 2583
 Db 2605 TTCAACACCATCATTTCTCCGTGATCTCGAAGACAGCTGACCCCTACGTCGCCACC 2664
 QY 2584 ATCAGCATCCACAGATATCAATTTTCGACCTCTGATCCTGATTAATTTGTTGAGC 2643
 Db 2665 ATCAGCCACCCATGACACCTCCATTTTTCGCGGTGATGAAGAAAGCCCTGGCAC 2724
 QY 2644 GATGGACGAGGAGGAATATCATGATGACACTCTGAGCGGATGGTATTAATGGAAG 2703
 Db 2725 GAGGCGAGCGGAGAG---AGCGCAAGCGCTGGGACAGCTGCGCTCCCGCCAGCAGCA 2781
 QY 2704 CACCCGAGCAGCCCTTCTATGAGTTTCACTTTCGAGGTTTCTTGTATGACATGGCTAC 2763
 Db 2782 CATCCAGACAGCCCTTCTATGAGTTTCACTTTCGCGGTTTCTGATGACAGCGCTAT 2841
 QY 2764 CCATATAATTATCCAAAGCCTATTGAGTATGAATACAT 2802
 Db 2842 CCCTCCGGTGCCGAGCGCTCAGAGCCGCGAGAGT 2880
 RESULT 12
 LOCUS AB023958 3460 bp mRNA linear ROD 09-MAR-2000
 DEFINITION Mus musculus MmLATS2 mRNA for 'warts/lats'-like kinase, complete cds.
 ACCESSION AB023958
 VERSION MmLATS2; warts/lats-like kinase.
 KEYWORDS MmLATS2; warts/lats-like kinase.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 Yabuta, N., Fujii, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A.,
 Nishiguchi, H., Endo, Y., Toji, S., Tanaka, H., Nishimune, Y. and
 Nojima, H.
 Structure, expression, and chromosome mapping of LATS2, a mammalian
 homologue of the Drosophila tumor suppressor gene lats/warts
 Genomics 63 (2), 263-270 (2000)
 20139436
 2 (bases 1 to 3460)
 Nojima, H. and Fujii, T.
 Direct Submission
 Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for
 Microbial Diseases, Osaka University, Department of Molecular
 Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
 (E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980,
 Fax:81-6-6875-5192)
 Sequence updated (06-Jan-2000).
 COMMENT Location/Qualifiers
 FEATURES
 source
 1..3460
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 1..3460
 /gene="MmLATS2"
 134..3262
 /gene="MmLATS2"
 /note="WALK"
 /codon_start=1
 /product="warts/lats-like kinase"
 /protein_id="BAA92380.1"
 /db_xref="GI:7212786"
 /translation="MRPKTFPATVSGNSRQLOETREGLKOPSKASTOGLLVGPNSD
 TSLDAKVLGSKDASROQMRKPKFGPYKALREINYSLLPANESCTSAAEVNRQM
 LAELVNAVGDCEMAGRALKPTSGRSTEALEYISKMGYLDPRNEQLVRVIKOTSPGK
 LAPTPVTRRPSFEGTEALPSYHRLGGANIEGPALEENPROYLDLFPFGAGTHGA
 QAHOHPKGYSTAVEPAHFPGTHYGRHLLSEOPGYGVORSSFNKPTPPDAYSSMA
 KAQGGPASILTPAHAGLYTASHHKPAATPGAHPLHLVLTGRGPTFTGESSAVALP
 SRNSLNADLYELGISTVPWSAAPLARRDSLQKQGLEASRPHVAFRAGPSRTNSFNNPQ

EBSLPAPNTVTAATAAHLHPVKSVRVLRPERPOTAVGSPHPAWAAPTAPATESLETK
 ESAGPHPLVDYGGSERRCPPPYPKHLILPSKSEQSVLDLSLCTSVQQLRGTE
 QDRSKSHKAGKAGRDKKQIOTSPVPRKNRDEKRESIRKISPYAFKPFMEQ
 HVENIKTYQKRSRRLQLEQAVAKALCEAEQOMRKILYOKESNRYLRKRMKDS
 MFVKITLIGAGFVCLACKLDTHALYAMKTLRKDLVLRNOVAHVAERDILAEAD
 NEWVKLYSFQDKSLYFVMDYIPGDDMSLLIRWEVPEHLARFYLAELTATESV
 HKMGFIHRDKFDNILDLDGHIKLTDLGCTGFRWTHNSKYQKGNHRQDSMEPGD
 LMDVSNCRGDRKLTLEORAKQHORLAHSLVGTPTNIAPEVLLRGYITQCCDWS
 VGVILFEMLVGPPFLAPTPTETOLKINWESTLHIPTPTQRLSABARDLITKLCAAD
 CRLGRDAGDLKAHPFNTIDFSRDIRKOPAPYVPTISHPMDTSPFPVDESPWHEA
 SGESAAKWDTLASPSKHPAHAFYEFTFRFRFDNDGYPFRCPKPEPAESADPGDADL
 EGAAEGCQPVVV"
 3440
 polyA_site
 BASE COUNT 861 a 992 c 963 g 644 t
 ORIGIN
 Query Match 25.1%; Score 806.2; DB 10; Length 3460;
 Best Local Similarity 64.9%; Pred. No. 8; e-184;
 Matches 1233; Conservative 0; Mismatches 648; Indels 18; Gaps 2;
 QY 919 CACTCTGCTAAATTTCTCAGCCTTCTGCCACTTACACTCAGCTCAGCTCAGCCGCTCCTATT 978
 Db 1304 CAACCTGAGCCCTCACTGCCGCCGCCCAACACGCTCACCGCTGACGCGCACATC 1363
 QY 979 CAACAGCCGCTGAAAGCATGCGCTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCCA 1038
 Db 1364 CTTCAACCTGTGAAGACGCTGCTGCTGCGGCCGAGCCCGACAGCCGTTGGGGCCC 1423
 QY 1039 ACCATCTCTTGTGATGCCACAGCCAGTTTACAGCTGTTTACGCTACCCCTTTTCTGAG 1098
 Db 1424 TCCACACCCGCTGGGTGGCTGCCGCCCAACAGCCTGCCACTGAGAGCTGGGAGAGCAAG 1483
 QY 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTCTGTTGCTGAACTCCAAAGCTATCAAGT 1158
 Db 1484 GAGGCGAGCGCCGCCACACCCGCTGGATGTGAGTATGCGGCTCCGAGCGCAGGTGC 1543
 QY 1159 CCACACCGCTTATCCAAACATCTGTACACCAAAACCCATCTGCTCCTCCTATATGAG 1218
 Db 1544 CCACCGCTCTCCGTACCAAGCACTTGTCTGCTGCCAGTAAGTCTGAGCAGTAGTACAGCGTG 1603
 QY 1219 TCAGTAAGTAAGCCCTGCCAAGATGAACAGCCCTGAGCTTACCCAAAGAGATGATAGTAG 1278
 Db 1604 GACCTGGACAGCTGTGTCACCACTGTGTCAGCAGAGTCTGCGAGGGGCGCTGAGCAAGAC 1663
 QY 1279 AAGAGTCCGAGACAGT-----GGTGACTCTGGGGATAAAGAAAAGAACAG 1323
 Db 1664 AGGAGTGACAAAGAGCACAAAGGTGCGAAGGGAGACAAAGCTGCCAGAGACAAAAAGCAG 1723
 QY 1324 ATTACAACTTCACTATCACTGTTCGGAAGAAACAGAAAGATGAAGAACGACGAGAGTCT 1383
 Db 1724 ATTACAGACTTCCCGGTGCTGTCGCCAGAGATAGCAGAGATGAAGAGAGAGAGTCT 1783
 QY 1384 CGGATTACAGAGTTACTCCCCACAGGCTTTAAGTCTTTCATGAGCAGCAGCTAGAGAAC 1443
 Db 1784 CGCATCAAGAGTTACTCCCTTTATGCTTCAATTTCTCATGAGCAGCAGCTGGAGAT 1843
 QY 1444 GTCCTGAAGTCTCATCAGCAGCGCTGTCATCGGAAGAACAGCTAGAAAATGAATGATG 1503
 Db 1844 GTCATCAAAACCTTACCAGCAGAGAGGTGAGCGGAGGTGAGCTGGAGAGGAGTGGCC 1903
 QY 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTTCCAGAGAGAG 1563
 Db 1904 AAAGCTGGGCTGTGTGAGCGGAGCAGGAGCAGATGAGGAAGATCTCTTACCAGAGAGG 1963
 QY 1564 TCTAACTATATTCGTCTTAAAGGGCTTAAATGGACAAGTCTATGTTTGTAAAGATAAAG 1623
 Db 1964 TCTAACTACAACCGCTGAAGAGGGCCCAAGATGCACAGTCCATGTTGTGAANAATCAAG 2023
 QY 1624 ACATTAGGAATAGGAGCGTTTGTGAAGTCTGTCTAGCAAGAAAAGTGTGATCTAAAGCT 1683
 Db 2024 ACTCTAGGCATCGTGCCTTTGGGGAAGTGTGCTGCTTGAAGCTGGACACACTCAGCT 2083

QY 1684 TTGTATGCAACAAAGACTCTTCGAAAGAAAGACGCTTGCTCCGAAATCAGGTGGCTCAT 1743
 DB 2084 CTGTACGCCATGAAGACTCTCAGGAAGAGGATCTCTGTAACCGAATCAAGTGGCCAT 2143
 QY 1744 GTGAAGCGGAGAGGGATATCTAGACAGAGCCACAAATGAGTGGTGGTCCGCTGTAC 1803
 DB 2144 GTCAAGGCTGAGAGGGACATCTCGCTGGAAGCAGACAAATGAGTGGTGGTCAAACTCTAC 2203
 QY 1804 TACTCTTTCAGGACAAAGCAACTCTGACTTTGTGATGAGTACTACATCTCTGGGGGGAT 1863
 DB 2204 TACTCTTTCAGGACAAAGCAGCTCTGACTTTGTGATGAGTACTACATCTCTGGGGGGAT 2263
 QY 1864 ATGATGAGCCTATTAATAGAAATGGGACCTTTCTCCGAAATCTGGCAGCATTTACATA- 1923
 DB 2264 ATGATGAGCCTGCTGATGAGTGGAGGCTCTTCCCTGAGCAGCTTGGCCGCTTCTACATT 2323
 QY 1924 GCAGAACTTACCTGTGAGTGAAGTGTTCATAAATGGGTTTATTCATAGAGATTT 1983
 DB 2324 GCAGAGTTGACCTGGCCATTTGAAGTGTCCACAAGATGGGCTTTATCCACCGGACATC- 2383
 QY 1984 AAACCTGATACATTTTATTGATGACCGTGGGACATTTAAATGACTGACTTTGGCTTG 2043
 DB 2384 AGCCTGACACATACATCTGACCTGATGGTGTCTATTAAGTGTGACAGATTTTGGCCTC 2443
 QY 2044 TGACCTGGCTTCAGATGACACATGACTCCAAGTACTACAGAGTGGGGATCAACCCACGG 2103
 DB 2444 TGCACTGGATTCAGTGGACTCAAAATTCAGAGTACTACAGAGTGGGGATCAACCCACGG 2503
 QY 2104 CAAGATAGCATGATTTTCACTAGCAATGGGAGATCTTCCAAATTTGGTGGGGAC 2163
 DB 2504 CAGGACAGCATGAGAGCGGTGACCTCTGGGACGATTTTCCAACTGCTGCTGGGAC 2563
 QY 2164 AGACTGAAGCCTGAGAGCGGAGCTGCTGCCAGCAGCAGCATCTAGGCCATTCT 2223
 DB 2564 AGTTAAGACCTGGAGAGAGGGCCAGAGCAACAGCAGAGTGGCTGGGACATCT 2623
 QY 2224 CTGTTGGGACTCCAAATTAATTTGACCTGAACTGAGTGTCTGCGAACAGGATATACACAG 2283
 DB 2624 CTGTTGGGACACCAAAATTAATTCAGCTGCGGAGTGTCTTCCGAAAGGATACACGAC 2683
 QY 2284 CTGTTGACCTGAGTGTGGTGTATTCTTTGTAATTTGGTGGGACACCTCTCT 2343
 DB 2684 CTGTTGACCTGAGTGTGGTGTATTCTTTGTAATTTGGTGGGACACCTCTCT 2743
 QY 2344 TTCTTGGCAACAAACCCATTTGAAGCAACAAATGAAGTGTATCTGCGCAAACTTCTCTA 2403
 DB 2744 TTCTTGGCCACCCACCCACAGAGACGAGTGTGAGTGTATCTGCGGAGACGACCTG 2803
 QY 2404 CACATCCCTCTCAAGCTAAGCTGAGTCTTGAAGCTCTGACCTCAATTAACACTGTGT 2463
 DB 2804 CATATCCCTACGAGGTGAGGCTCAGGCTGAGGCGGAGACCTCATCAGAGCTGTGC 2863
 QY 2464 CGAGGACCAAGAGCCCTCGCAAGACGGTGTCTGATGATAAGGCTCATCAATTT 2523
 DB 2864 TGGCGGCTGACTGCGCCCTGGGAGGATGGGCGAGATGACCTCAAGGACACCCCTTC 2923
 QY 2524 TTTAAGACCATCGATTTCTTGTGATGATCTGAGACAGAGTGTCTGCTATCATACCTCTAAA 2583
 DB 2924 TTCAACACCATCGATTTTCCCGTGATCTCGGAAGCAGCTGCACCTAGCTCCCAACC 2983
 QY 2584 ATCAGCATCCACAGATACATCAATTTGACCCCTTGTGATCTGATAAATTTGGAGC 2643
 DB 2984 ATCAGCACCCTGAGACACCTCAATTTGACCCGCTGGATGAAGAAAGCCCTGGCAC 3043
 QY 2644 GATGGCAGCAGGAGAAATATACATGACATCTGAGCGGATGGTATAAATGGGAAG 2703
 DB 3044 GAGGCCAGCGGAGAG---AGGCCAGCGCTGGGACAGCTGGGCTCCCCAGCAGCAG 3100
 QY 2704 CACCCCGCAGCAGCTTTCTATGATTTTCCACCTTTTGGAGGTTTTTGTGATGACAAATGGCTAC 2763
 DB 3101 CATCCAGACAGCCTTCTATGATGATTTTCCACCTTTCCGAGGTTTCTGATGACACAGCGCTAT 3160
 QY 2764 CCATATAATTTATCCAAAGCCTATTGAGTATGAATACATT 2802

DB 3161 CCCTTCGGTCCGCGAAGCCCTCAGAGCCGCGCAGAGGT 3199
 RESULT 13
 E38226
 LOCUS Human tumor regulatory gene.
 DEFINITION Human tumor regulatory gene.
 E38226
 ACCESSION E38226
 VERSION E38226.1 GI:18626935
 KEYWORDS JP 2000210086-A/1.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 5486)
 AUTHORS Koga,J., Kono,K. and N.Z.F.
 TITLE Human tumor regulatory gene
 JOURNAL Patent: JP 2000210086-A 1 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 COMMENT OS Homo sapiens (human)
 PN JP 2000210086-A/1
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
 PI C12N15/09, A61K31/00, A61K38/27, C07K14/47, C07K14/525, C12Q1/68, PC
 GO1N33/50,
 PC GO1N33/50, GO1N33/574, GO1N33/577//C12P21/08, (C12N15/09) PC
 C12R1:91), C12N15/00,
 CC A61K37/36, (C12N15/00, C12R1:91)
 PH Key Location/Qualifiers
 FT source 1.5486
 FT /organism="Homo sapiens"
 FT /db_xref="taxon:9606"
 BASE COUNT 1378 a 1452 c 1390 g 1266 t
 ORIGIN
 Query Match 24.5%; Score 787.2; DB 6; Length 5486;
 Best Local Similarity 68.7%; Pred. No. 3.5e-179;
 Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;
 QY 1220 CAGTAAGTACCCCTGCAAAAGATGAACAGCTTAGCTTACCCCAAGAGAGATGATAGTGA 1279
 DB 1999 CAGCATGGAGCAGAGCTCCGTGCGGGCCCAACAGAGCCGCGGCGGCAAGAGCC 2058
 QY 1280 AGATGCGGACAGTGGTACTCTGGGATAAAGAAAGAAAGAGATTAACACTTACCTA 1339
 DB 2059 GCAGAGCCGAGGGGCAAGAGCGGGAAGGATAAAGAGATTAAGAGATTCAGACTTCCCG 2118
 QY 1340 TCACTGTCGGAAGAAAGAGATGAAGAACGAGAGAGTCTTCGATTTCAGAGTTACT 1399
 DB 2119 TTCCGTCGCCAAACACAGCAGACGAGAGAGAGAGAGAGTCCAGCATCAAGAGTACT 2178
 QY 1400 CCCCACAGGCTTTAAGTCTTTCATGAGCAGCAGCTAGAGAGCTCTCTGAAGTCTCATC 1459
 DB 2179 CGCATACGCTTTAAGTCTTTCATGAGCAGCAGCTGGAGATGTCTCAAAACCTTACC 2238
 QY 1460 AGCAGCTCTGCAATCGGAAGAGCAGCTAGAAAATGAATGATGCGGGTTGGATTATCTC 1519
 DB 2239 AGCAGAGTTTACCGGAGGCTGCAGCTGGAGCAAGAAATGGCCAAAGCTTGGACTCTGTG 2298
 QY 1520 AAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAGAGATCTTAACATATATTCGTC 1579
 DB 2299 AAGCTGAGCAGGACAGATCGGAAGATCTCTTACCAAGAAAGAGTCTTAATTACAACAGGT 2358
 QY 1580 TTAAGAGGCTTAATGAGCAAGTCTATGTTTGAAGATAAAGACATTAAGTAATAGGAG 1639
 DB 2359 TAAAGAGGCCAAGATGGACAAGTCTATGTTTGTCAAGATCAAAACCTCGGGATCGGTG 2418


```
Db 2299 AAGCTGACGAGGACAGATCGCGAAGATCCTCTACCGAAGAGAGTCTAATTACACAGGT 2358
Qy 1580 TTAAGGCGCTAAATGACCAAGTCTATGTTTGAAGATTAAGACATTAGGAATAGGAG 1639
Db 2359 TAAGAGGCGCAAGTGGACCAAGTCTATGTTTGAAGATTAAGACATTAGGAATAGGAG 2418
Qy 1640 CGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCTGATCTAAAGCTTTTGTATCAACAAGAA 1699
Db 2419 CCTTTGGAGAAAGTGTCTTCTGTTGAAGTGGACACTCACGCCCTGTACGCCATGAAGA 2478
Qy 1700 CTCCTCGAAGAAAGACGCTTCTGCTCCGAATCAGTGGCTCATGTGAAGCGGAGAGGG 1759
Db 2479 CCCTTAAGAAAGAGATCTCTGCAACCGAATCAGTGGCCACCGCTCAAGGCCGAGAGGG 2538
Qy 1760 ATATCTAGCAGAGCCGACCAATGAGTGGTGGTCCGCTGTACTACTCTTTTCCAGGACA 1819
Db 2539 ACATCTGCGCGGACGACCAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGACA 2598
Qy 1820 AGGACAACCTTGTACTTTGTGATGGAATACATTCCTTGGGGGATATGATGAGCCTATTAA 1879
Db 2599 AAGACAGCCTGTACTTTGTGATGGAATACATTCCTTGGGGGATATGATGAGCCTATTAA 2658
Qy 1880 TTAGATGGGATCTTCTCTGAAATCTGGCAGCATCTTACATAGCAGCAATACCTCTGTG 1939
Db 2659 TCCGGATGGAGGTCTTCTCTGAGCAGCCTGGCCCGGTTCTACATCGCAGAGCTGACTTGG 2718
Qy 1940 CAGTTGAAGTGTTCATAAATGGTGTATTTATCTAGAGATATTAAACCTGATACATTT 1999
Db 2719 CCATTGAGAGTGTCCCAAGATGGCTTCATCCCGGAGACATCAAGCCTGATACATTT 2778
Qy 2000 TGATTGACCGTGTGGCCATATTAAATTTGACTGACTTTGGTCTGTGCACTGGCTTCAGAT 2059
Db 2779 TGATAGATCTGGATGTGCATTAATTAACCTACAGATTTTGGCCCTGTGCACTGGTTCAGT 2838
Qy 2060 GGACATGACTTCAAGTACTACAGAGTGGGATCAGCCAGCGCAAGATGATCATGATTT 2119
Db 2839 GGACTCAATATCCCAATATTACCAGAAAGGAGCCATGTGACAGCAGGACGATGGAGC 2898
Qy 2120 TCAGTAACCAATGGGAGATCCTTCCAATGTCGGTGTGGGACAGACTGAAGCCACTGG 2179
Db 2899 CCAGGACCTCTGGATGTGTCTACTTCTGCTGTGGGACAGGCTGAAGACCTTAG 2958
Qy 2180 AGCGGAGCTGTGCGCGCAGCAGCAGATGTCTAGGCCATCTCTGTGTTGGGACTCCCA 2239
Db 2959 AGCAGAGGCGCGGAGCAGCAGCAGAGGTGCTGGCAGCAATCTACTGTGGGACTCCA 3018
Qy 2240 ATTATATTGCACTGAAGTGTCTGCAAGCAGGATATACAGAGTGTGTGACTGGTGA 2299
Db 3019 ACTACATGCGACCGAGGTGCTCTCCGCAAGAGGTGACTCAACTCTGTGACTGGTGA 3078
Qy 2300 GTGTTGGTGTATTCTTTGTGAATGTTGGTGGGACACCTCTTCTTGGCACAAGCC 2359
Db 3079 GTGTTGGATGATCTCTTTCAGATGCTGTGGGCGCAGCCGCTTTTGGCACCCTACTC 3138
Qy 2360 CATTAGAACCAATGAAGTGTATCACTTGGCAAACTTCTTACACATCCCTCTCTCAAG 2419
Db 3139 CCACAGAAACCCAGCTGAAGTGTATCACTGGGAGAACACCTCCACATTCACGCCCAGG 3198
Qy 2420 CTAGCTGAGTCTCTGAAGCCTCTGACCTCATATCAACTGTGTGCGAGCACCAGAGACC 2479
Db 3199 TGAAGCTGAGCCCTGAGGCCAGGACCTCATCACCAAGCTGTGTGCTCCGACGACCACC 3258
Qy 2480 GCCTCGGCAAGACGCTCTGATGAGATGAAGGCTCATCTCAATTTTTTAAGACCATCGATT 2539
Db 3259 GCCTGGGCGGATGAGGCGCCATGACCTGAAGGCCACCCCTTCTTCAAGCCCATGACT 3318
Qy 2540 TCTCTAGTGTCTGAGACAGAGCTGTCTTCTTACATACCTCCCTAAATACCGCATCCAAAG 2599
Db 3319 TCTCCAGTGACATCCGGAAGCAGCAGCCAGCCCTAGTGTCCCAACCATCAGCCACCCATGG 3378
Qy 2600 ATACATCAATTTGACACCTGTGTGATCTGTATTAATTTGTGAGGATGGGACGAGGAGG 2659
Db 3379 ACACCTCAATTTGACACCCCGTAGATGAAGAGCCCTTGAAGAGTCCAGC---GAAG 3435
```

```
Oy 2660 AAAATATCAGTGACACTCTGAGCGGATGGTATAAAAATGGAAGCACCCGAGCAGCGTT 2719
Db 3436 GTAGCACCAAGCGCTGGGACACACTCACCTGCCCAATAACAAGCATCTGAGCAGCAT 3495
Qy 2720 TCTATGAGTTCACCTTCGAGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2779
Db 3496 TTTAGCAATTCACCTTCGGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3555
Qy 2780 AGCCTATTGATATCAATACATTTCATTCACAGGGCTCAGA 2819
Db 3556 AGCCTTCAGGACGACAGCTTCACAGGCTGAGAGCTCAGA 3595
```

RESULT 15
LOCUS AB028019 Homo sapiens HSLATS2 mRNA for large tumor suppressor 2, partial cds.
DEFINITION AB028019 3412 bp mRNA linear PRI 09-MAR-2000
ACCESSION AB028019
VERSION AB028019.1 GI:7212789
KEYWORDS large tumor suppressor 2; HsLats2.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Yabuta, N., Fujii, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Nishiguchi, H., Endo, Y., Toji, S., Tanaka, H., Nishimune, Y. and Nojima, H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
REFERENCE 2 (bases 1 to 3412),
AUTHORS Nojima, H. and Nishiguchi, H.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1999) Hiroshi Nojima, Research Institute for Microbial Diseases, Osaka University, Department of Molecular Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail: nojima@biken.osaka-u.ac.jp, tel:81-6-6875-3980, Fax:81-6-6875-5192)
FEATURES
source Location/Qualifiers
1. 3412
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 3412
/gene="HSLATS2"
/codon_start=3
/product="large tumor suppressor 2"
/protein_id="BA02381.1"
/db_xref="GI:7212790"

polyA_signal 3364 3369
BASE COUNT 801 a 1066 c 975 g 570 t
ORIGIN

Query Match 24.5%; Score 785.6; DB 9; Length 3412;
Best Local Similarity 68.6%; Pred No. 8.4e-179;
Matches 1098; Conservative 0; Mismatches 499; Indels 3; Gaps 1;

```
QY 1220 CAGTAAGTAAAGCCCTCAAGATGAACAGCTTAGCTTTACCCAGGAAGATGATAGTGAGA 1279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1489 CAGGCATGGAGCAGAGCCCTCGGTGCGGGCCCAACAGCAGCCGAGCGCGGACAGAGCC 1548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1280 AGAGTCGGACAGCTGTGACTCTGGGGATAAGAAAAGAAACAGATTCAACCTTCACCTA 1339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1549 GCAAAAGCCCAAGGGGGCAAAAGGGGAAAGGATAAAAGCAGATTTCAGACCTCTCCCG 1608
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1340 TCACCTGTCGGAATAACAAAGATGAAGAACGAAGAGTCTCGGATTCAGAGTTACT 1399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1609 TTCCTCCGCAAAACACAGACGACGAAGAGAGAGTACCCATCAAGAGCTACT 1668
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1400 CCCACAGCCCTTAAAGTCTTCATGGAGCAGCAGCTAGAGACGCTCTGGAAGTCTCATC 1459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1669 CGCATACGCCCTTAAAGTCTTCATGGAGCAGCAGCTGGAGATGTATCAAAACCTACC 1728
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1460 AGCAGCTCTGCATCGGAAGAACAGCAGTACAAATCAATGATGCGGGTTGGATTATCTC 1519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1729 AGCAGAAGGTAAACCGGAGCTCGACCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTG 1788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1520 AAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTTCGTC 1579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1789 AAGCTGACGAGGAGCAGATCGGGAAGATCCTCTACCAAGAGAGTCTAATTACAACAGGT 1848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1580 TTAAGGGCTAAATGGCAAGCTATGTTTCTAAAGATAAGACATTTAGGAATAGGAG 1639
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1849 TAAAGAGGGCAAGATGGCAAGCTATGTTTCTCAAGATCAAAACCCCTGGGGATCGGTG 1908
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1640 CGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCTGATATCAAGCTTTGTATGCAACAAAGA 1699
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1909 CCTTTGGAGAAAGTGTGCTTGTCTGTTGTAAGTGGACACTCAGCCCTGTAGCCCATGA 1968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1700 CTCTTGGAAAGAAAGAGTTCTCTCGGAATCAGTGGCTCATGTGAACGGGAGAGG 1759
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1969 CCCTAAGGAAAGAGGTCTCTGAAACGGAAATCAGTGGCCCAAGCTCAAGGCCGAGGG 2028
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1760 ATATCTAGCAGAGCGGACAAATGAGTGGTGTGCGCCTGTACTCTTTCCAGGACA 1819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2029 ACATCTGGCCGAGGAGACAAATGAGTGGTGTGCTCAACTCTACTACTCTTCCAAGACA 2088
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1820 AGGACAACTTGTACTTTGTGATGGACTACATCTCTGGGGGATATGATGAGCCTATFAA 1879
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2089 AAGACAGCCTGTACTTTGTGATGGACTACATCTCTGGTGGGACATGATGAGCCTGTGA 2148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1880 TTAGATGGGCATCTTTCCTGAAAATCTGCGACATCTACATAGCAGAACTTACCTGTG 1939
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2149 TCGGATGGAGGTCTTCCCTTGAGCACCCTGGCCGGTCTACATCGCAGAGCTGACTTTG 2208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1940 CAGTTGAAGCTGTCAATAAATGGTTTTATTCATAGATATTAACCTGTATAACATTT 1999
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2209 CCATTTGAGAGTGTCCCAAGATGGGCTTATCCACCAGACATCAAGCCTGTATAACATTT 2268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2000 TGATTTACCTGTATGGCCATATTAAATTTGACTGACTTTGGCTTTGTGCACTTGGCTTCAGAT 2059
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2269 TGATAGATCTGGATGGTCACTATAAATCAAGATTTCGGCCCTCTGCACTTGGGTTCAGGT 2328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2060 GGACACATGACTCCAAGTACTACAGAGTGGGATCACCACCGCAGATAGCATGATT 2119
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2329 GGACTCAAAATTCAAAATATTTACCAAGAGGGAGCCCATGTTCAGACAGGACGATGGAGC 2388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2120 TCAGTAACGAATGGGAGATCCCTTCCAAATGTGCGTGTGGGGACAGACTGAAGCCACTGG 2179
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2389 CCACGACCTCTGGATGATGTGTCTAACTGTGCTGCTGCGTGTGGGGACAGGCTGAAGACCCTAG 2448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: January 16, 2003, 18:29:01
Job time : 6446.88 secs

```
QY 2180 AGCGGAGAGCTGCTCGCCAGCACCAGCAGATGCTTAGCCCAATCTCTGTTGGGACTCCCA 2239
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2449 AGCAGAGGCGCGGAAGCAGCAGCAGAGTCTGCTGGCACATTCACCTGTTGGGACTCCAA 2508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2240 ATTATATTGCACCTTGAAGTGTCTACTGCGGAACAGGATATACACAGCTGTGTGACTGGTGA 2299
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2509 ACTACATCGACCCGAGGTGCTCTCCGCAAAAGGGTACACTCAACTCTGTGACTGTGGA 2568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2300 GTGTTGGTCTTATCTTCTGTAATGCTGTGGGACAACTCTCTTTTGGGCACAAACC 2359
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2569 GTGTTGGAGTGATCTCTTCGAGATGCTGTGGGCGGCGCCCTTTTGGGACCTACTC 2628
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2360 CATTAGAAACACAAATGAAGTTATCATCTGGCAAACTTCTTACACATCTCCCTCTCAAG 2419
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2629 CCACAGAAACCCAGCTGAAGGTGATCACTGGGAGAACACGCTCCACATTCAGGCCAGG 2688
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2420 CTAAGTGTAGTCTCTGAAGCTCTGACCTCATTTATCAAACTGTGTGAGGACCAAGAGACC 2479
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2689 TGAAGCTGAGCCTTGAGCCAGGACCTCATCACCAGCTGTGCTCTCCGAGACCAACC 2748
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2480 GCCTCGGCAAGACGCTGCTGATGAGATAAGGCTCATCATTTTAAAGACCATCGATT 2539
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2749 GCCTGGGCGGAATGGGCGCGATGACCTGAGGCCACCCCTTCTTCAGGCCATTTGACT 2808
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2540 TCTCTAGTGTATCTGAGACAGCAGTCTGTTCATATACATCCCTAAAATCAGCATCCCAACAG 2599
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2809 TCTCCAGTGACATCGGGAAGCAGCAGCCCTACCTTCCACCATCAGCCACCCCATGG 2868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2600 ATACATCCAAATTCGACCCCTGTGATCTCTGTATTAATTTGGAGCGATGCGGAGGAGG 2659
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2869 ACACCTCGAATTCGACCCCTGTAGATGAAGAAGCCCTTGGAAACGATGCCAGC 2925
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2660 AAAATATCAGTGACACTCTGAGCGGATGGTATAAAAATGGGAAGCACCCCGAGCAGCTT 2719
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2926 GTAGCACCAGGCTGGACACACTCACCTCGCCCAATAACAAGCATCTGAGCAGCAT 2985
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2720 TCTATGAGTTACCTTTCGGAGGTTTTTGTATGACAAATGGCTACCCATATATATCCAA 2779
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2986 TTTACGAATTCACCTTCCGAAGGTTCTTTGTATGACAATGGCTACCCCTTTCGATGCCCAA 3045
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2780 AGCCTATTGAGTATGAATACATTTCACAGGGCTCAGA 2819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3046 AGCCTCAGGACGAGAAGCTTCACAGGCTGAGAGCTCAGA 3085
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 409.821 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: US-09-763-334-3

Perfect score: 3213

Sequence: 1 gtgcaacattcaattacgg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3213	100.0	3213	21	AAZ51506
2	3209.8	99.9	3213	17	AAZ51506
3	2483	77.3	3984	17	AAZ42119
4	2483	77.3	3984	21	AAZ42118
5	2481.4	77.2	7382	20	AAZ51505
6	1946.8	60.6	2442	20	AAZ87396
7	1126.8	35.1	1374	20	AAZ32981
8	1012	31.5	1357	22	AAZ32982
9	812.6	25.3	3155	17	AAZ42120

10	812.6	25.3	3155	21	AAZ51507
11	787.2	24.5	5276	20	AAZ51508
12	787.2	24.5	5486	21	AAZ51509
13	787.2	24.5	5486	21	AAZ51510
14	785.6	24.5	1961	21	AAZ61160
15	785.6	24.5	3533	22	AAZ17165
16	730.6	22.7	1912	23	ABK43457
17	713.8	22.2	1498	21	AAZ61158
18	681	21.2	2043	22	AAZ03989
19	598.2	18.6	795	20	AAZ32984
20	521.2	16.2	5720	17	AAZ42117
21	521.2	16.2	5720	21	AAZ51508
22	519.6	16.2	3319	23	ABL03169
23	518.6	16.1	638	21	AAZ79942
24	492.4	15.3	1501	22	AAZ75341
25	468.4	14.6	582	20	AAZ32983
26	376.6	11.7	678	21	AAZ79982
27	371.8	11.6	676	22	AAZ71179
28	371.8	11.6	676	23	ABK43776
29	252.4	7.9	11187	23	ABL03168
30	248.4	7.7	1935	20	AAZ06834
31	248.4	7.7	4983	22	AAZ44629
32	244.4	7.6	3018	17	AAZ31452
33	244.4	7.6	3018	24	ABK84773
34	244.4	7.6	3800	22	ABA08668
35	226.4	7.0	1710	21	AAZ43403
36	226.4	7.0	2003	21	AAZ49841
37	221.6	6.9	1894	21	AAZ39567
38	220	6.8	1922	21	AAZ49426
39	211.6	6.6	2001	21	AAZ36174
40	210.2	6.5	2255	23	ABL29695
41	198.6	6.2	2101	17	AAZ31451
42	179.8	5.6	2160	22	AAZ66703
43	175.6	5.5	734	23	ABK43709
44	175.2	5.5	734	22	AAZ27161
45	175.2	5.5	734	22	AAZ56728

ALIGNMENTS

RESULT 1
AAZ51506
ID AAZ51506 standard; DNA: 3213 BP.

XX AC AAZ51506;

XX DT 21-JUN-2000 (first entry)

XX DE Mouse Lats (large tumour suppressor) DNA.

XX KW Mouse: Lats; large tumour suppressor; cytostatic; vulnerary;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX OS Mus musculus.

XX FH Key Location/Qualifiers

XX CDS 1..2889
FT /**tag= a
FT /product= "Lats protein"
FT /partial

XX WO200010602-A1.

XX PD 02-MAR-2000.

XX XX 18-AUG-1999; 99WO-US19068.

XX PF

Db 1561 GAGTCTAATATATCGCTTTAAAGGGCTAAATGGACAAAGTCTATGTTTGAAGATA 1620
 Qy 1621 AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGTGATACTAAA 1680
 Db 1621 AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGTGATACTAAA 1680
 Qy 1681 GCTTTGATGCAACAAAGACTCTTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGGCT 1740
 Db 1681 GCTTTGATGCAACAAAGACTCTTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGGCT 1740
 Qy 1741 CATCTGAAGCGGAGAGGATATCTTACGACAAAGCGGACATGAGTGGTGGTCCGCGTG 1800
 Db 1741 CATCTGAAGCGGAGAGGATATCTTACGACAAAGCGGACATGAGTGGTGGTCCGCGTG 1800
 Qy 1801 TACTACTCTTTCCAGGACAAAGCAACTTGTACTTTGTGTGGACTACATTCCTGGGGGG 1860
 Db 1801 TACTACTCTTTCCAGGACAAAGCAACTTGTACTTTGTGTGGACTACATTCCTGGGGGG 1860
 Qy 1861 GATATGATGAGCCATTAAATPAGAAATGGGCATCTTTCTGAAATCTGGCAGATTCTAC 1920
 Db 1861 GATATGATGAGCCATTAAATPAGAAATGGGCATCTTTCTGAAATCTGGCAGATTCTAC 1920
 Qy 1921 ATACGACAACTTACTGTGACGTTGAAAGTGTTCATAAATGGGTTTATTCATAGAGAT 1980
 Db 1921 ATACGACAACTTACTGTGACGTTGAAAGTGTTCATAAATGGGTTTATTCATAGAGAT 1980
 Qy 1981 ATTAACCTGTAAACATTTTCAATGACCGTGTATGCGCATATTAATTTGACTGCTTGGC 2040
 Db 1981 ATTAACCTGTAAACATTTTCAATGACCGTGTATGCGCATATTAATTTGACTGCTTGGC 2040
 Qy 2041 TTGTGCACTGGCTTCCAGATGACACATGACTCCAAAGTACTACCAAGTGGGGATCACCCA 2100
 Db 2041 TTGTGCACTGGCTTCCAGATGACACATGACTCCAAAGTACTACCAAGTGGGGATCACCCA 2100
 Qy 2101 CGGCAAGATAGCATGGATTTTCAATGACCGTGTATGCGCATATTAATTTGACTGCTTGGC 2160
 Db 2101 CGGCAAGATAGCATGGATTTTCAATGACCGTGTATGCGCATATTAATTTGACTGCTTGGC 2160
 Qy 2161 GACAGACTGAAGCCACTGGAGCGAGAGTCTGCTCGCAGCAGCAGCATGCTAGCCCAT 2220
 Db 2161 GACAGACTGAAGCCACTGGAGCGAGAGTCTGCTCGCAGCAGCAGCATGCTAGCCCAT 2220
 Qy 2221 TCTCTGTTGGGACTCCCAATTAATTTGACCTGAGTGTCTGCGAACAGGATATACA 2280
 Db 2221 TCTCTGTTGGGACTCCCAATTAATTTGACCTGAGTGTCTGCGAACAGGATATACA 2280
 Qy 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTCTTTGTGAAATGTTGGTGGGACACCT 2340
 Db 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTCTTTGTGAAATGTTGGTGGGACACCT 2340
 Qy 2341 CTTTCTTGGCACAACCCCAATTAAGAACACAAATGAAGTGTATCATCTGGCAAACTTCT 2400
 Db 2341 CTTTCTTGGCACAACCCCAATTAAGAACACAAATGAAGTGTATCATCTGGCAAACTTCT 2400
 Qy 2401 CTACACATCCCTCTCAAGCTTAAGCTGAGTCTGAGTCTGAGGCGCTGACCTCATTAACAATG 2460
 Db 2401 CTACACATCCCTCTCAAGCTTAAGCTGAGTCTGAGTCTGAGGCGCTGACCTCATTAACAATG 2460
 Qy 2461 TGTCGAGGACCAAGAGCCGCTCGGCAAGAACGGTGTGATGAGATAAAGGCTCATCCA 2520
 Db 2461 TGTCGAGGACCAAGAGCCGCTCGGCAAGAACGGTGTGATGAGATAAAGGCTCATCCA 2520
 Qy 2521 TTTTAAAGACCATCGATTTCTCTAGTGTATCTGAGACGAGTCTGCTTCAATACATCCCT 2580
 Db 2521 TTTTAAAGACCATCGATTTCTCTAGTGTATCTGAGACGAGTCTGCTTCAATACATCCCT 2580
 Qy 2581 AAAATCAGCATCCAAAGATACATCCAAATTTTCGACCTGTGTGATCTGATAAATTTGG 2640
 Db 2581 AAAATCAGCATCCAAAGATACATCCAAATTTTCGACCTGTGTGATCTGATAAATTTGG 2640
 Qy 2641 AGCGATGCGCAGGAGGAGAAATATCAGTGACACTCTGACGCGATGATATAAATTTGG 2700
 Db 2641 AGCGATGCGCAGGAGGAGAAATATCAGTGACACTCTGACGCGATGATATAAATTTGG 2700

Qy 2701 AAGCACCCGAGCAGCGCTTTCTATGAGTTCACCTTTTCGAGGTTTTTTTGTGACAAATGGC 2760
 Db 2701 AAGCACCCGAGCAGCGCTTTCTATGAGTTCACCTTTTCGAGGTTTTTTTGTGACAAATGGC 2760
 Qy 2761 TACCCATATAATTAATCCAAAGCCTATTGAGTATGAATACATTCATTCACAGGGCTCAGAA 2820
 Db 2761 TACCCATATAATTAATCCAAAGCCTATTGAGTATGAATACATTCATTCACAGGGCTCAGAA 2820
 Qy 2821 CAACAGTCTGATGAGATGATCAACACACAGCTCCGATGAAACAAACCGAGATCTAGTG 2880
 Db 2821 CAACAGTCTGATGAGATGATCAACACACAGCTCCGATGAAACAAACCGAGATCTAGTG 2880
 Qy 2881 TATGTTTAAATAACTAGGAGATCATTTGTAAGAATTTGCAAGAGGCTTGAAGTGCAGGGGT 2940
 Db 2881 TATGTTTAAATAACTAGGAGATCATTTGTAAGAATTTGCAAGAGGCTTGAAGTGCAGGGGT 2940
 Qy 2941 TTTTGAAGTTTTGAGAAAAATATGCAAAATGACAGAGTTTTGTGCTCTGTGTACAATA 3000
 Db 2941 TTTTGAAGTTTTGAGAAAAATATGCAAAATGACAGAGTTTTGTGCTCTGTGTACAATA 3000
 Qy 3001 TTTTATTTTCTTAAGTTATGGAAATTTGTTTAAATGCTTAAATTTATCCACCCTTTAA 3060
 Db 3001 TTTTATTTTCTTAAGTTATGGAAATTTGTTTAAATGCTTAAATTTATCCACCCTTTAA 3060
 Qy 3061 TTCAGTAAATTTAGAAAAATTTGTTTAAAGAAAGTAAATTTGAACCTGAGTATTATATGTC 3120
 Db 3061 TTCAGTAAATTTAGAAAAATTTGTTTAAAGAAAGTAAATTTGAACCTGAGTATTATATGTC 3120
 Qy 3121 AATCTTGGTAACTTAAAGTACTTAAAGAGAGAGCCTGGTATCTTTTGTATATATAATAA 3180
 Db 3121 AATCTTGGTAACTTAAAGTACTTAAAGAGAGAGCCTGGTATCTTTTGTATATATAATAA 3180
 Qy 3181 ATAATTTTAAATCCCAAAAAAATAAAAAA 3213
 Db 3181 ATAATTTTAAATCCCAAAAAAATAAAAAA 3213

RESULT 2
 AAT42119
 ID AAT42119 standard; cDNA; 3213 BP.
 XX AC AAT42119;
 XX AC AAT42119;
 DT 31-JAN-1997 (first entry)
 XX M-lats gene encoding large tumour suppressor.
 DE Mouse; m-lats gene; large tumour suppressor; fetal brain;
 KW protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnary; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX Mus musculus.
 OS Mus musculus.
 XX OS
 XX OS
 FH Key Location/Qualifiers
 CD 1..2889
 FT /*tag= a
 FT /product= m-lats protein
 XX
 XX
 XX WO9630402-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 26-MAR-1996; 96WO-US04101.
 XX
 PR 27-MAR-1995; 95US-0411111.
 XX
 XX (UYUA) UNIV YALE.
 XX
 XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 PI
 XX

DR WPI: 1996-455275/45.
 DR P-PSDB; AA05179.
 XX New isolated large tumour suppressor gene - used to develop prods:
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Disclosure; Page 126-130; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats protein,
 CC and has been isolated from a newborn mouse brain phage lambda-TAP
 CC cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
 CC (AA742117). A homologous mouse sequence has also been isolated
 CC (m-lats2, AAT42120). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;
 Query Match 99.9%; Score 3209.8; DB 17; Length 3213;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3211; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGGTTCTTAAGAGTCTCTAGTTCT 60
 DB 1 GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGGTTCTTAAGAGTCTCTAGTTCT 60

QY 61 CAGACAGCGGCCCATCTCTAGAGAAATGTGTTTATCGTTCTGAAAGCCCAACTCA 120
 DB 61 CAGACAGCGGCCCATCTCTAGAGAAATGTGTTTATCGTTCTGAAAGCCCAACTCA 120

QY 121 CAGCGGATAGGAAGACCTCTGTGATCGCGCATTTGCATCAAGCTCAC 180
 DB 121 CAGCGGATAGGAAGACCTCTGTGATCGCGCATTTGCATCAAGCTCAC 180

QY 181 CCAAGCAATGGACAGAGTGAACCCCAACCACTTCAAGTTAGGAGTGTACTCTCT 240
 DB 181 CCAAGCAATGGACAGAGTGAACCCCAACCACTTCAAGTTAGGAGTGTACTCTCT 240

QY 241 CCACCACTCCGAGAGCCAGACCCACCTCCCGAGGACCACTCCCGCTCCCGCTCA 300
 DB 241 CCACCACTCCGAGAGCCAGACCCACCTCCCGAGGACCACTCCCGCTCCCGCTCA 300

QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCGA 360
 DB 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCGA 360

QY 361 ATCTCCCTCTTCCACTGGGGGTGCGAGAGGGTACCCCTCCACCACTCTTACCACT 420
 DB 361 ATCTCCCTCTTCCACTGGGGGTGCGAGAGGGTACCCCTCCACCACTCTTACCACT 420

QY 421 TCTCCCATGAATCCCTTAGCCAGGCTCAGAGGCGCATTAAGTTCTGTTCAGTTGGTAGA 480
 DB 421 TCTCCCATGAATCCCTTAGCCAGGCTCAGAGGCGCATTAAGTTCTGTTCAGTTGGTAGA 480

QY 481 CAACCCATCATATGACAGTACTAGCAAAATTTAACTTTACACCGGGCGACCTGGAGTT 540
 DB 481 CAACCCATCATATGACAGTACTAGCAAAATTTAACTTTACACCGGGCGACCTGGAGTT 540

QY 541 CAGAATGGTGTGCTAGTCTGATTTTATCGTCAACCAAAATGTCCCACTGGTTCTGTG 600
 DB 541 CAGAATGGTGTGCTAGTCTGATTTTATCGTCAACCAAAATGTCCCACTGGTTCTGTG 600

QY 601 ACTCGGAGCCACCACTCTCTATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660
 DB 601 ACTCGGAGCCACCACTCTCTATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660

QY 661 TTACAAACAGGGGCTTGTGCTGCCACCATCATTCGCCAATGGAACGTTCTCTCAGTCG 720
 DB 661 TTACAAACAGGGGCTTGTGCTGCCACCATCATTCGCCAATGGAACGTTCTCTCAGTCG 720

QY 721 ATGATGGTCCCAACAGCAAGCATTAACATGGAGCTTTATAATATTAAATGTCCTGGA 780
 DB 721 ATGATGGTCCCAACAGCAAGCATTAACATGGAGCTTTATAATATTAAATGTCCTGGA 780

QY 781 CTGAAACAGCGCTGCCCGCAGTCTGCTCTCTGCGCAGTCTATCCCAAGCGGTGG 840
 DB 781 CTGAAACAGCGCTGCCCGCAGTCTGCTCTCTGCGCAGTCTATCCCAAGCGGTGG 840

QY 841 CATGAAATTCCTACATGCAACCTTAACATACAGTGAAGTCAAAATCTTTTAATAACCCA 900
 DB 841 CATGAAATTCCTACATGCAACCTTAACATACAGTGAAGTCAAAATCTTTTAATAACCCA 900

QY 901 TTAGCAAGTAGAGCAAGTCACTCTGCTAATCTCAGCCTTCTGCCACTACAGTCACTGCC 960
 DB 901 TTAGCAAGTAGAGCAAGTCACTCTGCTAATCTCAGCCTTCTGCCACTACAGTCACTGCC 960

QY 961 ATCACACCGCGTCTCTTATTAACAGCCCGTGAAGACATGCGCGTCTGAAACCGAGCTG 1020
 DB 961 ATCACACCGCGTCTCTTATTAACAGCCCGTGAAGACATGCGCGTCTGAAACCGAGCTG 1020

QY 1021 CAGACTGCTTTAGCCCCCAACCCATCTCTTGGATGCCACAGCCAGTTCAGACTGTTCCAG 1080
 DB 1021 CAGACTGCTTTAGCCCCCAACCCATCTCTTGGATGCCACAGCCAGTTCAGACTGTTCCAG 1080

QY 1081 CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCTGTCTATCCACCTGTTGCTGAA 1140
 DB 1081 CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCTGTCTATCCACCTGTTGCTGAA 1140

QY 1141 GTCCTCAAGTATCAAGTCCACCGCTTATCCAAACATCTGCTACACCAACCA 1200
 DB 1141 GTCCTCAAGTATCAAGTCCACCGCTTATCCAAACATCTGCTACACCAACCA 1200

QY 1201 TCTGTCCCTCCATATGAGTCAAGTGAAGCCCTGCAAGATGAACAGCTAGCTTACCC 1260
 DB 1201 TCTGTCCCTCCATATGAGTCAAGTGAAGCCCTGCAAGATGAACAGCTAGCTTACCC 1260

QY 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGACTCTGGGGATTAAGAAAGAA 1320
 DB 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGACTCTGGGGATTAAGAAAGAA 1320

QY 1321 CAGATTACAATTCACCTATCTGTCGGAAGATGAAGAGATGAAGAGAGAGAG 1380
 DB 1321 CAGATTACAATTCACCTATCTGTCGGAAGATGAAGAGATGAAGAGAGAGAGAG 1380

QY 1381 TCTCGGATTTCAGATTCTCCACAGGCTTTAAGTTCTTCTATGAGCAGCAGCTAGAG 1440
 DB 1381 TCTCGGATTTCAGATTCTCCACAGGCTTTAAGTTCTTCTATGAGCAGCAGCTAGAG 1440

QY 1441 AAGCTCCTGAAGTCTCATCAGAGCTCTGCTGGAAGAGAGCTAGAAATGAATG 1500
 DB 1441 AAGCTCCTGAAGTCTCATCAGAGCTCTGCTGGAAGAGAGCTAGAAATGAATG 1500

QY 1501 ATGCGGGTGGATTATCTCAAGTCCAGGATCAATGAGAAAGATGCTTTGCCAGAAA 1560
 DB 1501 ATGCGGGTGGATTATCTCAAGTCCAGGATCAATGAGAAAGATGCTTTGCCAGAAA 1560

QY 1561 GAGTCTAACTATATTCGTCTTAAAGGGCTTAAATGAGCAAGTCTATGTTGTAAGATA 1620
 DB 1561 GAGTCTAACTATATTCGTCTTAAAGGGCTTAAATGAGCAAGTCTATGTTGTAAGATA 1620

QY 1621 AAGACATATGGAATAGAGGCTTGGTGAAGTCTGTAGCAGAAAGCTGATCTATAA 1680
 DB 1621 AAGACATATGGAATAGAGGCTTGGTGAAGTCTGTAGCAGAAAGCTGATCTATAA 1680

QY 1681 GCTTTCTATGCAACAAGACTCTTCAAGAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCT 1740
 DB 1681 GCTTTCTATGCAACAAGACTCTTCAAGAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCT 1740

QY 1741 CATGTGAAGCGGAGAGGATATCTTAGCAGAAAGCCGACAAATGAGTGGGTGCTCGCGCTG 1800

Db 1741 CATGTGAAGCGGAGGAGATCTCTAGCAGAAGCGCAATGAGTGGTGGTCCGCGTG 1800
 Qy 1801 TACTACTCTTTCCAGGACAGGACAACTGTACTTTGTGATGGACTACATTCCTGGGGGG 1860
 Db 1801 TACTACTCTTTCCAGGACAGGACAACTGTACTTTGTGATGGACTACATTCCTGGGGGG 1860
 Qy 1861 GATATGATGAGCCTATTAAATAGATGGCATCTTCTGAAATCTGGCAGATTCAC 1920
 Db 1861 GATATGATGAGCCTATTAAATAGATGGCATCTTCTGAAATCTGGCAGATTCAC 1920
 Qy 1921 ATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTTTATTATAGAGAT 1980
 Db 1921 ATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAATGGGTTTATTATAGAGAT 1980
 Qy 1981 ATTAACCTGATACATTTTCAATGACCGTGTATGCCATATTAATTCAGTACCTTTGGC 2040
 Db 1981 ATTAACCTGATACATTTTCAATGACCGTGTATGCCATATTAATTCAGTACCTTTGGC 2040
 Qy 2041 TTGTGCACTGGCTTCAGATGACACATGACTCCAAAGTACTACAGAGTGGGATCACCA 2100
 Db 2041 TTGTGCACTGGCTTCAGATGACACATGACTCCAAAGTACTACAGAGTGGGATCACCA 2100
 Qy 2101 CGCAAGATAGCATGATTTTCAATGACCGTGTATGCCATATTAATTCAGTACCTTTGGC 2160
 Db 2101 CGCAAGATAGCATGATTTTCAATGACCGTGTATGCCATATTAATTCAGTACCTTTGGC 2160
 Qy 2161 GACAGACTGAAGCCACTGGAGGAGAGCTGCTGCCAGCAGCAGGATGTCTAGCCCAT 2220
 Db 2161 GACAGACTGAAGCCACTGGAGGAGAGCTGCTGCCAGCAGCAGGATGTCTAGCCCAT 2220
 Qy 2221 TCTCTGTTGGGACTGCCCAATTAATTCACCTGAAGTGTCTGCAAGAGGATATACA 2280
 Db 2221 TCTCTGTTGGGACTGCCCAATTAATTCACCTGAAGTGTCTGCAAGAGGATATACA 2280
 Qy 2281 CAGCTGTGTGACTGGTGGAGTGTGGTGTATTTCTGTAAGTGTGGTGGGACAACT 2340
 Db 2281 CAGCTGTGTGACTGGTGGAGTGTGGTGTATTTCTGTAAGTGTGGTGGGACAACT 2340
 Qy 2341 CCTTCTTGGCACAACCCCAATTAAGAACACAAATGAGGTATCATCTGCAAACTTCT 2400
 Db 2341 CCTTCTTGGCACAACCCCAATTAAGAACACAAATGAGGTATCATCTGCAAACTTCT 2400
 Qy 2401 CTACATCTCTCTCAAGCTAAGTGTAGTCTGAGTCTGAGCCTCTGACCTCATTAACAATG 2460
 Db 2401 CTACATCTCTCTCAAGCTAAGTGTAGTCTGAGTCTGAGCCTCTGACCTCATTAACAATG 2460
 Qy 2461 TGTGAGGACGACGAGACCGCTCGGCAAGACGGTGTCTGATGAGATAAGGCTCATCCA 2520
 Db 2461 TGTGAGGACGACGAGACCGCTCGGCAAGACGGTGTCTGATGAGATAAGGCTCATCCA 2520
 Qy 2521 TTTTGTAGACCATGATTTCTCTAGTGTATCTGATGATCTGACAGCAGTCTGCTATACATCCCT 2580
 Db 2521 TTTTGTAGACCATGATTTCTCTAGTGTATCTGATGATCTGACAGCAGTCTGCTATACATCCCT 2580
 Qy 2581 AAAATCAGCATCCCAACAGATACATCCCAATTTTCGACCTGTTGATCTGATAAATGTGG 2640
 Db 2581 AAAATCAGCATCCCAACAGATACATCCCAATTTTCGACCTGTTGATCTGATAAATGTGG 2640
 Qy 2641 AGCGATGCGAGGAGGAGAAATATCAGTACACTCTGAGCGGATGTTATATAAATGGG 2700
 Db 2641 AGCGATGCGAGGAGGAGAAATATCAGTACACTCTGAGCGGATGTTATATAAATGGG 2700
 Qy 2701 AAGCACCCGAGCAGCCTTTCTATGATGTTACCTTTCCGAGGTTTTCATGACAAATGGC 2760
 Db 2701 AAGCACCCGAGCAGCCTTTCTATGATGTTACCTTTCCGAGGTTTTCATGACAAATGGC 2760
 Qy 2761 TACCATATTAATATCCAAAGCCTATTGATGATGAATACATTCATTCAGGCTCAGAA 2820
 Db 2761 TACCATATTAATATCCAAAGCCTATTGATGATGAATACATTCATTCAGGCTCAGAA 2820
 Qy 2821 CAACAGTCTGATGAAGATGATCAACACAAAGCTCCGATGGAAACCAACCGATCTAGTG 2880

Db 2821 CAACAGTCTGATGAAGATGATCAACACACAAAGCTCCGATGGAAACACACCGAGATCTAGTG 2880
 Qy 2881 TATGTTTAAATAAAGCTAGGAGATCATTTGAAGATTTTGAAGAGGCTCAAGTGCAGGGGT 2940
 Db 2881 TATGTTTAAATAAAGCTAGGAGATCATTTGAAGATTTTGAAGAGGCTCAAGTGCAGGGGT 2940
 Qy 2941 TTTTGAAGCTTTTGAGAAATTTATGCAAAATGTGACAGAGTCTGTGCTCTGTGTACAATA 3000
 Db 2941 TTTTGAAGCTTTTGAGAAATTTATGCAAAATGTGACAGAGTCTGTGCTCTGTGTACAATA 3000
 Qy 3001 TTTTATTTTCTTAAGTATGGAATTTCTTTTAAATGTTTAAATGTTTATTTATCCACCTTTTAA 3060
 Db 3001 TTTTATTTTCTTAAGTATGGAATTTCTTTTAAATGTTTAAATGTTTATTTATCCACCTTTTAA 3060
 Qy 3061 TTTTATTTTCTTAAGTATGGAATTTCTTTTAAATGTTTAAATGTTTATTTATCCACCTTTTAA 3120
 Db 3061 TTTTATTTTCTTAAGTATGGAATTTCTTTTAAATGTTTAAATGTTTATTTATCCACCTTTTAA 3120
 Qy 3121 AATTTCTGCTACTTAAAGTACTTTAAAGAGAGAGGCTGCTATCTTTTGTATATATAATAA 3180
 Db 3121 AATTTCTGCTACTTAAAGTACTTTAAAGAGAGAGGCTGCTATCTTTTGTATATATAATAA 3180
 Qy 3181 ATAAATTTAAATCCCAAAAAA 3213
 Db 3181 ATAAATTTAAATCCCAAAAAA 3213

RESULT 3
 AAT42118
 ID AAT42118 standard; cDNA; 3984 BP.
 XX
 AC AAT42118;
 XX
 DT 31-JAN-1997 (first entry)
 XX

XX H-lats gene encoding large tumour suppressor.

Human; h-lats gene; large tumour suppressor; fetal brain;
 protein-serine/threonine-kinase; cell proliferation; antisense;
 dominant-negative; cancer; degenerative disorder; trauma;
 growth deficiency; therapy; antitumour; vulnerability; diagnostic;
 transgenic plant; transgenic animal; growth; senescence; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
 XX CDS 231..3623
 XX FT /*tag= a
 XX FT /product= Lats protein

XX W09630402-A1.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-041111.

XX (UYA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;

XX WPI: 1996-455275/45.

XX P-PSDB: AAW05178.

XX New isolated large tumour suppressor gene - used to develop prods.
 XX for inhibiting cell proliferation or for enhancing proliferation

XX Claim 6; Page 118-123; 215pp; English.

XX This sequence encodes a human large tumour suppressor h-lats protein,
 CC and has been isolated from a fetal human brain phage lambda-gt10.
 CC cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene

CC (AAT42119). The full-length cDNA is present in plasmid pBS(KS)-h-lats.
 CC The gene encodes a putative protein-serine/threonine-kinase, and
 CC inhibits cell proliferation and plays a crucial role throughout
 CC development. Activators or inhibitors of lats function (e.g. an
 CC antisense oligonucleotide or dominant-negative lats fragment) may be
 CC used in therapy of cancer or other proliferative disorders,
 CC degenerative disorders, trauma, growth deficiency, etc., and
 CC fragments of the gene may be used as diagnostic probes. A
 CC lats-inhibitor sequence may be expressed in a transgenic plant or
 CC farm animal to confer increased growth and inhibit senescence.
 XX
 SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match 77.3%; Score 2483; DB 17; Length 3984;

Best Local Similarity 87.4%; Pred. No. 0;

Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;

QY 1 GTCCACATCAATTAACCCAAACAAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCTCT 60
 DB 732 GTCCAGCAATCACTTAACCCCAACAGAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCTCT 791
 QY 61 CAGACAGACGGCCCATCTCTAGGAGAAATGTGTTTATCTGTTCTGAAAGCCCAACTCA 120
 DB 792 CAGAGCATGGCCGCCCATCTAGGAGAAAGTGTGSCCTATCATTTCTGAGAGTCCCAACTCA 851
 QY 121 CAGGCGGATGTAGAGACCTCTGTCTGGATCCGGCATTCGACATTTGCTCAAGTCTAC 180
 DB 852 CAGACAGATGTAGGAAGACCTTTGTCTGGATCTGGTATATACAGCATTTGTTCAAGTCTAC 911
 QY 181 CCAAGCAATGGACAGAGATGAACCCCAACACCACTCAAGTTAGGAGTGTACTCTCT 240
 DB 912 CTTAGCAAGGACAGAGATGAACCCCAACCACTCAAGTTAGGAGTGTACTCTCT 971
 QY 241 CCACCACTCCGAGAGCCAGACCCCACTCCCGAGGACCACTCCCGCTCCCGCTCA 300
 DB 972 CCACCACTCCGAGAGCCAGACTCCCGTCCAGAGGTACACTCCACTCCCGCTCA 1031
 QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCCGA 360
 DB 1032 TGGGAACCAAACTCTCAACAAAGCGCTATCTGGAACATGGAATACGTAATCTCCCGA 1091
 QY 361 ATCTCCCTCTTCCACTCGGGGCTGGCAGGAGGCTACCTCCACCACTCTTACCACT 420
 DB 1092 ATCTCTCTCTGTCGACCTGGGACATGCAAGAGGCTATCTCTCCACCACTCTCAACACT 1151
 QY 421 TCTCCCATGAATCCCTCATGACAGGCTCAGAGGCGCATTAAGTTCTCTGTCAGTTGTTAGA 480
 DB 1152 TCCCCCATGAATCCCTCATCAAGACAGAGAGGCTATGTTCTCTGTTGGCAGA 1211
 QY 481 CAACCCATCATGACAGAGTACTAGCAATTTAACTTTACACCGGGGCGAGCTGGAGTT 540
 DB 1212 CAACCAATCATGACAGAGTTCTAGCAAAATTTAACTTTCCATCAGGAGAGCTGGAATG 1271
 QY 541 CAGAATGTTGGTGGTCACTGATTTTATCGTCCACCAAAA---TGTCCTCCACTGGTTCT 597
 DB 1272 CAGAATGTTGGTGGTCACTGATTTTATCGTCCACCAAAA---TGTCCTCCACTGGTTCT 1331
 QY 598 GTGACTCGGCGAGCCACCTCCATATCTCTGACCCCACTAATGGACAAAGCCCTCT 657
 DB 1332 GTGAATCGGCGAGCCACCTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTCT 1391
 QY 658 GCTTTTCAACAGGGGCTTCTGCTGTCACCACTATTCGCCAATGGAAGCTTCTCTAG 717
 DB 1392 GCTTTTCAACAGGGGCTTCTGCTGTCACCACTATTCGCCAATGGAAGTATTCCTCTAG 1451
 QY 718 TCCATGATGTTGCCCAACAGGACAGTCAATACATGAGCTTTATATATTAATGTCCT 777
 DB 1452 TCTATGATGTTGCCCAACAGGAAATAGTCAATACATGGAATATATAGTGTACCT 1511
 QY 778 GGAATGCAACAGCCCTGGGCCCCAGTCTCTGCTCTGCGCAGTCACTCCCAAGCGT 837
 DB 1512 GGAATGCAACAAATTTGGCTCAGTCTCTGCTCTGCTCCAGCCAGTCACTCCCGGACAGT 1571

QY 838 GGCATGAATAATTCCTACATGGCAACCTTAACATACCAGTGAAGTCAAAATCTTTTAATAAC 897
 DB 1572 GGCATGAATAATTCCTACATGGCAACCTTAACATACCAGTGAAGTCAAAATCTTTTAATAAC 1631
 QY 898 CATTAGGAGTAGAGCAAGTCACTCTGCTAAATCTCAGCTCTTCCACTACAGTCACT 957
 DB 1632 CATTAGGAGTAGAGCAAGTCACTCTGCTAAATCTCAGCTCTTCCACTACAGTCACT 1691
 QY 958 GCATCACACCCCTCTCTATTCAACAGCCGCTGAAAGCATGCGCTCTGAAACCCAGAG 1017
 DB 1692 GCAATTACACCACTCTCTATTCAACAGCCGCTGAAAGCATGCTGTATTAAACCCAGAG 1751
 QY 1018 CTCGACAGTCTTTAGCCCAACCCATCTCTTTGGATGCCAGCAGCTAGACTTT 1077
 DB 1752 CTACAGAGTCTTTAGCCCAACCCATCTCTTTGGATGCCAGCAGCTAGACTTT 1811
 QY 1078 CAGCTTACCCCTTTTCTGAGGCTACAGCTTCAAGTGTGCTCTCATCCACCTGTTGCT 1137
 DB 1812 CAACCCAGTCTTTTCTGAGGAGCCGCTTCAATGTGACTGTGATGCCACCTGTTGCT 1871
 QY 1138 GAAGCTCCAAAGCTTATCAAGGCTCCACCCCTTATCCAAACATCTGCTTACCAAAAC 1197
 DB 1872 GAAGCTCCAAACCTTATCAAGGCTCCACCCCTTATCCAAACATCTGCTTACCAAAAC 1931
 QY 1198 CATCTGCTCCCTCCATATGAGTCAAGTAAAGTAAAGCCCTGCAAGATGAACAGCTTACGTTA 1257
 DB 1932 CATCTGTTCTCCATACAGGCTCAATCAGTAAAGCTAGCAAGAGGATGACCCAGCTT 1991
 QY 1258 CCAAGSAGATGATAGTGAAGAGTGGGACAGTGGTCACTCTGGGGATAAGAAAG 1317
 DB 1992 CCAAGSAGATGAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 2051
 QY 1318 AAACAGATTACAACTTCACTCTGTTGCGAAACAAAGAAAGATGAAGAACGAAGA 1377
 DB 2052 AAACAGATTACAACTTCACTCTGTTGCGAAACAAAGAAAGATGAAGAACGAAGA 2111
 QY 1378 GAGTCTGGATTCAGAGTACTCCCAAGGCTTTAAGTCTTCTGAGAGAGCAGCTA 1437
 DB 2112 GAATCTGATTTCAAAAGTATTTCTCTCAAGCATTTAAATTTCTTTTGGAGCAACTGA 2171
 QY 1438 GAGAAGCTCTGAAGTCTCATCAGAGGCTCTCATCGGAAGAGCAGCTAGAAAATGA 1497
 DB 2172 GAAGATGACTCAAACTCATCAGAGGCTCTCATCGGAAGAGCAGCTAGAAAATGA 2231
 QY 1498 ATGATCGGGTTGGATTTATCTCAAGATGCCAGGATCAATGAAGAGATGCTTTGCCAG 1557
 DB 2232 ATGATCGGGTTGGATTTATCTCAAGATGCCAGGATCAATGAAGAGATGCTTTGCCAG 2291
 QY 1558 AAAGAGTCTAACTTATTCGTCTTAAAGGGCTTAAATGGACAGTCTATGTTGTAAG 1617
 DB 2292 AAAGAGTCTAAATTTACATCGCTTAAAGGGCTTAAATGGACAGTCTATGTTGTAAG 2351
 QY 1618 ATAAAGCATTTAGGAATAGGAGGTTTGGTGAAGTCTCTAGCAAGAAAGTCTGATACT 1677
 DB 2352 ATAAAGCATTTAGGAATAGGAGGTTTGGTGAAGTCTCTAGCAAGAAAGTCTGATACT 2411
 QY 1678 AAAGCTTTGTATGCAACAAAGACTTCTCGAAAGAAAGAGCTTCTGCTCGAAATCAGGTG 1737
 DB 2412 AAGGCTTTGTATGCAACAAAGACTTCTCGAAAGAAAGAGTCTTCTCTCGAAATCAGGTG 2471
 QY 1738 GCTCATGTGAAGCGGAGAGGATATCTTAGCAGAGCCCAATAGTGGGTGGTCCGC 1797
 DB 2472 GCTCATGTGAAGCGGAGAGGATATCTTAGCAGAGCCCAATAGTGGGTGGTGGTTCGT 2531
 QY 1798 CTGACTACTCTTTCCAGGACAAAGGACAACTTGTACTTTGTGATGAGCTACATTTCTGGG 1857
 DB 2532 CTATATTTATTTCCAGGACAAAGGACAAATTTATCTTTGTAATGAGCTACATTTCTGGG 2591
 QY 1858 GGGATATGATGAGCTTATTAATGAAGGCGCATCTTCTCTGAAATCTGGCAGCATTC 1917
 DB 2592 GGTGATATGATGAGCTTATTAATGAAGGCGCATCTTCTCTGAAATCTGGCAGCATTC 2651
 QY 1918 TACATAGCAGAACTTACCTGTCAGTGTGAAGTGTTCATAAATGGTGTATTATCATAGA 1977

Db 2652 TACATAGCAGAACTTTACCTGTGCAAGTGTGAAGAGTGTTCATAAAATGGTTTATTATCATAGA 2711
QY 1978 GATATTAACCTGATAACATTTTGTGATGACCGTGTGATGCGCATATTAATAATTGACTGACTTT 2037
Db 2712 GATATTAACCTGATAATTTTGTGATGATGCTGTGATGCTGATGCTGATGCTGATGCTGATGCT 2771
QY 2038 GCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2097
Db 2772 GGCCTCTGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2831
QY 2098 CCACGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157
Db 2832 CCACGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2891
QY 2158 GGGGACAGACTGAACCCACTGGAGCGGAGAGCTGCTGCCAGCACCAGCGATGCTAGCC 2217
Db 2892 GGAGACAGACTGAACCCACTGAAGCGGAGAGCTGACCCAGCACCAGCGATGCTAGCA 2951
QY 2218 CATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2277
Db 2952 CATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3011
QY 2278 ACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2337
Db 3012 ACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3071
QY 2338 CTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2397
Db 3072 CTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3131
QY 2398 TCTCTACACATCCCTCCCTCAAGCTAAGCTGAGTCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCT 2457
Db 3132 TCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3191
QY 2458 CTGCT 2517
Db 3192 CTTTCCGAGGACCCGAGAGATCGCTTAGCGAAGATGCTGCTGATGATGATGATGATGATGATGAT 3251
QY 2518 CCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2577
Db 3252 CCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3311
QY 2578 CTTAAATTCACACATCCACAGATACATCCATTTTCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 2637
Db 3312 CTTAAATTCACACATCCACAGATACATCCATTTTCGACCTGCTGCTGCTGCTGCTGCTGCTGCT 3371
QY 2638 TGGAGCGATGCGAGCGAGGAGAAATATCAGTACACTCTGAGCGGATGGTATGATGATGATGAT 2697
Db 3372 TGGAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3431
QY 2698 GGGAGGACCCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2757
Db 3432 GGAAGAGCTGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3491
QY 2758 GCTACCCATATTAATTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2817
Db 3492 GCTACCCATATTAATTCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3551
QY 2818 GAACACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2877
Db 3552 GAGCAGCAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3611
QY 2878 GTGATGCTTAAATTAACCTGAGGATCATTTGTA ---GAATTCGACAGGCGCTGAAGTG 2933
Db 3612 GTATATGCTTAAACACCTAGTAAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 3671
QY 2934 CAGGGCTTTTGAAGCTTTTGTAG ---AAAATATGCAAAATGACAGAG ---TTTGTGT 2985
Db 3672 CGAGGTGTTTGTAGGCTTCTGAGAGTAAATATGCAATATGACAGAGCTATATATGATGATGAT 3731
QY 2986 GCTCTGTGTACAAATATTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3045

Db 3732 GCTCTGTGTACAAATATTTTATTTTCTCTAAATATATGGGAAATCCTTTTAAATGTTAATTT 3791
QY 3046 ATTCACACCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3105
Db 3792 ATTCACACCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3851
QY 3106 CTGAGTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3154
Db 3852 CTGATATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3900

RESULT 4

AAZ51505
ID AAZ51505 standard; DNA; 3984 bp.
XX
AC AAZ51505;
XX
DT 21-JUN-2000 (first entry)
XX
Human Lats (large tumour suppressor) DNA.
XX
Human; Lats; large tumour suppressor; cytostatic; vulnery;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; knock-out mouse; ds.
OS Homo sapiens.
XX
Key Location/Qualifiers
FH 231...3623
FT CDS
FT /*tag= a
FT /product= "Lats protein"

WO200010602-A1.

02-MAR-2000.

18-AUG-1999; 99WO-US19068.

18-AUG-1998; 98US-0096996.

18-AUG-1998; 98US-0096997.

(UYVA) UNIV YALE.

Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
Trenchalk GS, Stewart RA;

WPI: 2000-246496/21.

P-PSDB: AAY70390.

Use of lats proteins, complexes of lats and cdc2 for treating cancer
that is refractory to treatment by standard chemotherapy and radiation
therapy, and disorders associated with aberrant levels of cdc2 activity

Claim 44; Fig 12; 134pp; English.

The present sequence is a DNA encoding human Lats (large tumour
suppressor) protein which is a cell overproliferation inhibitor and a
negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
The present sequence is useful for treating cancer that is refractory
to standard chemotherapy or radiation therapy such as hyperplasia,
metaplasia, or dysplasia, and disorders associated with aberrant
levels of cdc2 activity. Conditions treated by promoting cdc2 function
include degenerative disorders, growth deficiencies, hypoproliferative
disorders, physical trauma, lesions, and wounds. An animal model
preferably a mouse, in which a lats gene has been disrupted by homologous
recombination, e.g. a lats knock-out mouse, is used for screening
compounds that can be used to treat or prevent cancer, particularly
skin cancer, soft tissue sarcomas and ovarian tumours, and disorders

CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
XX
SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match	77.3%;	Score 2483;	DB 21;	Length 3984;
Best Local Similarity	87.4%;	Pred. No. 0;		
Matches 2769;	Conservative	0;	Mismatches 385;	Indels 15; Gaps 4;
QY	1	GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGAGTTCTTAAGAGTCTCTAGTTCCT 60		
DB	732	GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGAGTTCTTAAGAGTCTCTAGTTCCT 791		
QY	61	CAGAGACGCGCCATCTCTAGAGAAATGTGTTATCGTTCTCAAGAGCCCACTCA 120		
DB	792	CAGAGGATGCGCCGACCTAGGAGAAAGTGTGCGCTCATCTCTGAGAGTCCCACTCA 951		
QY	121	CAGCGGATGTAGAGACCTCTGTCTGGATCCGGATTCGACGATTTGCTCAAGCTCAC 180		
DB	852	CAGACAGATGTAGAGACCTTTGTCTGGATCTGGTATATATATATATATATATATAT 911		
QY	181	CCAAAGCAATGGACAGAGTGAACCCCGCCACACCTCAAGTTAGGAGTGTACTCCT 240		
DB	912	CCTAGCAACGACAGAGTGAACCCCGCCACACCTCAAGTTAGGAGTGTACTCCT 971		
QY	241	CCACCACCTCGAGAGCCGACCCACCTCCCGAGGACCACTCCCGCTCCCGCTCA 300		
DB	972	CCACCACCTCCAGAGGCGGACGACTCCCGCTCCAGAGGTACACTCCACCTCCCGCTCA 1031		
QY	301	TGGGAACCAAGCTCTCAGACAAAGCGTACTCTGGGAACATGGAGTACGTAATCTCCGA 360		
DB	1032	TGGGAACCAACTCTCAACAAAGCGTATTCTTGAAACATGGAATACGTAATCTCCGA 1091		
QY	361	ATCTCCCTGTTCACCTGGGCGTGGCAGGAGGGTACCTCCACCACTCTTACACT 420		
DB	1092	ATCTCTCTCTCCACCTGGGCGTGGCAGGAGGGTATCTCCACCACTCTTACACT 1151		
QY	421	TCTCCATGAATCCCTTAGCAGGCTCAGAGGGCCATTTAGTTCTGTCTCCAGTTGGTGA 480		
DB	1152	TCCCGCATGAATCCCTTAATCAAGGACAGAGAGGCATTAGTTCTGTCTGTGGCAGA 1211		
QY	481	CAACCATCATCATGACAGTACTAGCAATTTAATTTACACAGCGGACCTGGAGTT 540		
DB	1212	CAACCAATCATCATGACAGTACTAGCAATTTAATTTTCCATCAGGAGACCTGGAGT 1271		
QY	541	CAGATGTGTGGTGGTCACTCTGATTTTATCGTCAACAAAAA---TGTCCTCCACTGGTCT 597		
DB	1272	CAGATGTGTGGTGGTCACTCTGATTTTATCGTCAACAAAAA---TGTCCTCCACTGGTCT 1331		
QY	598	GTGACTGGCAGCCACCACTCCATATCTCTGACCCAGCTATGGACAAGCCCTCT 657		
DB	1332	GTGAATGGCAGCCACCACTCCATATCTCTGACAGCAGTAAATGGACAAGCCCTCT 1391		
QY	658	GCTTTACAAACAGGGCTTCTGCTCTCCACATCATCTCGCAATGGAAAGCTTCTCTCAG 717		
DB	1392	GCTTTACAAACAGGGGATCTGCTCTCTCTCTCTATATACAAATGGAAATTTCTCTCAG 1451		
QY	718	TCGATGATGGTCCCAACAGGAACAGTATACATGGAGCTTTAATATTAATGTCCCT 777		
DB	1452	TCTATGATGGTCCCAACAGGAATATGATACATGGAATATATATATATATATATAT 1511		
QY	778	GGAGTCAAAACAGCTGGGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 837		
DB	1512	GGAGTCAAAACAAATTTGGGCTCAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1571		
QY	838	GGCATGAAATCTTACATGGCAACCTTAACATACCAGTGGAGTCAAAATCTTTTAATAC 897		
DB	1572	GGCATGAAATCTTACATGGCAACCTTAACATACCAGTGGAGTCAAAATCTTTTAATAC 1631		
QY	898	CCATTAGGAAGTAGACAAAGTCACTCTGCTTAATCTCGACCTTCTGCGCACTACACTCACT 957		
DB	1632	CCATTAGGAAGTAGACAAAGTCACTCTGCTTAATCTCGACCTTCTGCTACAAAGTCACT 1691		

QY	958	GCCATACACACCGCTCTCTATTTCAACAGCCCTGAAAGCATCGCGTCTCTCAACACGAG 1017		
DB	1692	GCAATTACACAGCTCTCTATTTCAACAGCCCTGAAAGTAGTATGCTGTTATTAACACGAG 1751		
QY	1018	CTGACAGCTCTTTAGCCCAACCCATCTCTTGGATGCCACAGCCAGTTTCAGACTGTT 1077		
DB	1752	CTACAGACTGCTTTAGCACCTTACACACCTCTTCTGGATACCACAGCCATTTCAACTGTT 1811		
QY	1078	CAGCTTACCCCTTTTCTGAGGCTACAGTCAAGTGTGCTGTCTATCCACCTCTGTTGCT 1137		
DB	1812	CAACCCAGCTCTTTCTGAGGACCCCTCAATGTGACTGTGATGCCACCTGTTGCT 1871		
QY	1138	GAAGTCCCAAGCTATCAAGGTCCACACCGCTTATCCAAACATCTGCTACACCAAC 1197		
DB	1872	GAAGTCCCAACTATCAAGGACCCACACCTTACCCAAACATCTGCTGCACCAAC 1931		
QY	1198	CCATCTGCTCCCTCATATGAGTCAAGTAAAGCTTCAAGCTTCAAGAGTGAACAGCTAGCTTA 1257		
DB	1932	CCATCTGTTCTCTCCATACGATCAATGATAGCTTCAAGCTTCAAGAGGATCAGCAAGCTTG 1991		
QY	1258	CCCAAGGAAGATGATGAGTGAAGAGTGGGACAGTGGTGTGCTCTGGGGATTAAGAAAG 1317		
DB	1992	CCCAAGGAAGATGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 2051		
QY	1318	AAACAGATTACAACTTCACTATCACTCTCGAAGAAACAAAGAAAGATGAAGAAAGAA 1377		
DB	2052	AAACAGATTACAACTTCACTATCACTCTCGAAGAAACAAAGAAAGATGAAGAAAGAA 2111		
QY	1378	GAGTCTCGGATTCAAGATTCTCCACAGGCTTTAAAGTTCTTCAATGAGGAGCAGCAGTA 1437		
DB	2112	GATCTCTGTTTCAAGTTATTTCTCTCAAGCATTTAAATTTCTTTTATGAGCAACATGTA 2171		
QY	1438	GAGAACTCTGAAAGTCTCATCAGAGCTTGCATCGAAGAAAGCAGCTACAAATGAA 1497		
DB	2172	GAAATGTACTCAAACTCTCATCAGAGCTTACATCTGTAAGAAACAAATGAGAGTAA 2231		
QY	1498	ATGATCGGCTGGATTTCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAG 1557		
DB	2232	ATGATCGGCTGGATTTCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAG 2291		
QY	1558	AAAGCTTAACTATATCTCTTAAAGGGCTTAAATGGACAGCTATGTTTGAAG 1617		
DB	2292	AAAGATCTAATFACATCTCTTAAAGGGCTTAAATGGACAGCTATGTTTGAAG 2351		
QY	1618	ATAAGACATTAGAAATAGGAGCTTTGGTGAAGTCTGCTAGCAAGAAAGTCTGATCT 1677		
DB	2352	ATAAGACATTAGAAATAGGAGCTTTGGTGAAGTCTGCTAGCAAGAAAGTCTGATCT 2411		
QY	1678	AAAGCTTTGTATCAACAAAGCTTCTCGAAGAAAGAGCTTCTGCTCGAAATCAGGTG 1737		
DB	2412	AAGCTTTGTATCAACAAAGCTTCTCGAAGAAAGAGCTTCTGCTCGAAATCAGGTG 2471		
QY	1738	GCTCATGTGAAGGGAGAGGATATCTTAGCAGAGCCGACATGAGTGGGTGGTCCGC 1797		
DB	2472	GCTCATGTGAAGGGAGAGGATATCTTAGCAGAGCTTCTGCTCGAAATCAGGTG 2531		
QY	1798	CTGTACTACTCTTTCAGGACAGGACACTTGTACTTTGTGATGAGTACATCTCTGGG 1857		
DB	2532	CTATATTATTTCAATCCCAAGATAAGGACAAATTTATCTTTGTAATGAGTACATCTCTGGG 2591		
QY	1858	GGGATATGATGAGCCCTTAAATTAGAATGGCATCTTCTCTGAAATCTGGCAGCATTC 1917		
DB	2592	GGTATATGATGAGCCCTTAAATTAGAATGGCATCTTCTCTGAAATCTGGCAGCATTC 2651		
QY	1918	TACATACAGAACTTACCTGTGCAAGTGAAGTGTTCATATAATGGGTTTTTATTCATAGA 1977		
DB	2652	TACATACAGAACTTACCTGTGCAAGTGAAGTGTTCATATAATGGGTTTTTATTCATAGA 2711		
QY	1978	GATATTAAACCTGATACATTTTGATGAGCGGATGGCCATATTAAATTCAGTACTGCTT 2037		
DB	2712	GATATTAAACCTGATACATTTTGATGAGCGGATGGCCATATTAAATTCAGTACTGCTT 2771		
QY	2038	GCCTTGCTGCTGCTTTCAGATGACACATGACTCCCAAGTACTACCAAGTGGGGATCAC 2097		

Db	2772	GGCCTCTGCACATGGCTTCAGATGGACACACAGATTCTAAGTACTATCAAGATGGTGGACCAT	2833
Qy	2098	CCACGGCAAGATAGCATGGATTTCAGTAACGAATGGGAGAGATCCTTCCAAATTTGTCGGTGT	2157
Db	2832	CCACGGCAAGATAGCATGGATTTCAGTAATGNAATGGGGGATCCTCAAGCTGTCATGT	2891
Qy	2158	GGGACAGACTGAAGCCACTGGAGCGGAGAGTGTGCCACAGCACAGCGATGTCATAGCC	2217
Db	2892	GGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCACAGCACAGCGATGTCATAGCA	2951
Qy	2218	CATTCTCTGGTTGGGACTCCCAATTAATATGCACCTGAAGTGTACTGCCAACAGGATAT	2277
Db	2952	CATTCTCTGGTTGGGACTCCCAATTAATATGCACCTGAAGTGTGTCTACCAACAGGATAC	3011
Qy	2278	ACACAGCTGTGTGACTTGGTGGAGTCTGGTGTATTCTTCTGGAATGTTGGTGGGACAA	2337
Db	3012	ACACAGTGTGTGATTTGGTGGAGTCTGGTGTATTCTTCTTGGAAATGTTGGTGGGACAA	3071
Qy	2338	CTCTCTTTCTTGGCACAAACCCCAATTAGAAACACAAATGAAGTTATCATCTGCGCAAACT	2397
Db	3072	CTCTCTTTCTTGGCACAAACCCCAATTAGAAACACAAATGAAGTTATCAACTGGCAAAACA	3131
Qy	2398	TCCTCTACACATCCCTCTCTCAAGCTAAGCTGAGTCTGCTGAAGCCTCTGACCTCATATCAAA	2457
Db	3132	TCCTCTTCACATTCCACACACAGCTAAGCTCAGTCTGCTGAAGCTTCTGATCTTATTTATAA	3191
Qy	2458	CTGTGTCGAGGACCAAGACACCGCTCGGCAAGACGGTGTGATGAGATAAAGGCTCAT	2517
Db	3192	CTTTGCCGAGGACCGGAGATCGTTAGCGAAGAAATGGTGTGATGAATAAAGCTCAT	3251
Qy	2518	CCATTTTTTAAGACCATCGATTTCTCTAGTGTATGTGACAGCAGCTGCTGCTTCATACATC	2577
Db	3252	CCATTTTTTAAGACCATTTGACTCTCCAGTGCACCTGACCTGACAGCAGCTGCTGCTTCATACAT	3311
Qy	2578	CTTAAATCACGCATCCACAGATACATCCAAATTTGACCCCTGTTGATCCTGATAAATTG	2637
Db	3312	CCTTAAATCACACACCCACAGATACATCAAAATTTGATCTGTTGATGCTGATGAATAA	3371
Qy	2638	TGGAGCGATTGGCAGGAGGAGGAAAATATCAGTCACACTCTGAGCGGATGGTATAAAAAT	2697
Db	3372	TGGAGTGATGATACGAGGAGAAAATGTAATGNACACTCTCAATGATGGTATAAAAAT	3431
Qy	2698	GGGAGCACCCCGAGCAGCGTTTCTATGAGTTCACCTTCCGAGGTTTTTTGATGACAAT	2757
Db	3432	GGAAAGCATCTCGAATCATCTATCAATTTACCTTCCGAAGGTTTTTTGATGACAAT	3491
Qy	2758	GGCTACCCATATAATATCCAAAGCCTATTGAGTATGNAATACATTCATTACACAGGCTCA	2817
Db	3492	GGCTACCCATATAATATCCGAAGCCTATTGAATATGAATACATTAATTCACAGGCTCA	3551
Qy	2818	GAACAACAGCTGTGATGAAGATGATCAACACACAAGCTCCGATGCAACACACCGAGATCTA	2877
Db	3552	GAGCACAGTCGGATGAAGATGATCAAAACAGCTCAGAGATTTAAAATCGCGATCTA	3611
Qy	2878	GTGATATGTTTAATAACTAGGAGATCATTTGTA - - - - -GAAATTTGCAAGAGCGCTGAAGTG	2933
Db	3612	GTATATGTTTAACACACATAGTAATAATATGTAATCAGGATTTGTAAAGGGCTGAAATG	3671
Qy	2934	CAGGGGTTTTGAAGTTTTGTAG - - - - -AAAAATATGCAAAATGTGACAGAG - - - - -TTTGTGT	2985
Db	3672	CGAGGTGTTTTGAGGTTCTGAGAGTAAAAATTTATGCAAAATATGACAGAGCTATATATGCTG	3731
Qy	2986	GCCTCTGTGACAAATATTTTATTTTCTTCCCTAAGTTATGGGAAATGTTTTAAAATGTTAATTT	3045
Db	3732	GCCTCTGTGACAAATATTTTATTTTCTTCCCTAATTTATGGGAAATGTTTTAAAATGTTAATTT	3791
Qy	3046	ATTCCACCCCTTTTAATTCAGTAATTTAGAAAAAATTTGTTATGAAGAAAGTAAATTTATGAA	3105
Db	3792	ATTCCACCGCTTTAATTCAGTATTTAGAAAAAATTTGTTATGAAGAAAGTAAATTTATGAA	3851
Qy	3106	CTGAGTATTATAGTCAATTTCTTGGTACTTAAAGTACTTAAAGAGGAGAG 3154	

3952 CTGAATATTATAGTCAGCTCTTGCTACTTAAAGTACTTAAAAATAAGTAG 3900

RESULT 5
AAX87396
ID AAX87396 standard; cDNA; 7382 BP.
XX AC
XX AAX87396;
XX 08-OCT-1999 (first entry)
XX DT
XX Human WART1 cDNA.
XX DE
XX WART1; hWART1; WART orthologue; human; signal transduction;
XX KW protein kinase; cancer; tumour; diagnosis; therapy; ss.
XX KW
XX OS Homo sapiens.
XX OS
FH Key Location/Qualifiers
FH 5'UTR 1..399
FT FT /*tag= a
FT FT 12..63
FT FT /*tag= b
FT FT /*note= "Contains 10 copies of GGC repeat, similar
FT FT to repeats that undergo expansion in human
FT FT diseases associated with neuronal
FT FT phenotypes"
CDS 395..3787
FT FT /*tag= c
FT FT 3788..7382
FT FT /*tag= d
FT FT 6058..6346
FT FT repeat_region
FT FT /*tag= e
FT FT /*note= "Alu-J subfamily repeat motif"
FT FT polyA_signal 7339..7344
FT FT /*tag= f
FT FT replace(978,6)
FT FT /*tag= g
FT FT /*note= "Results in Ala/Gly change"
FT FT variation replace(1840)
FT FT /*tag= h
FT FT /*note= "Silent polymorphism"
FT FT variation replace(3252..3253)
FT FT /*tag= i
FT FT /*note= "deletion of 2 adenines, causes
FT FT frameshift"
XX XX
XX WO9937787-A2.
XX XX
XX 29-JUL-1999.
XX XX
XX 20-JAN-1999; 99WO-US01145.
XX PF
XX 21-JAN-1998; 98US-0072023.
XX PR
XX XX
XX (SUGE-) SUGEN INC.
XX PA
XX Flanagan P, Plowman GD;
XX PI
XX WP1; 1999-458698/38.
XX DR
XX P-PSDB; AAY06526.
XX XX
XX New nucleic acid encoding human orthologs of Drosophila WART
XX PT proteins; used to identify specific modulators for treating cancer
XX PT or for diagnosis
XX XX
XX Claim 1; Page 120-121; 137pp; English.
XX XX
XX This is the nucleotide sequence of a cDNA clone coding for a human
XX CC orthologue, i.e. hWART1 (see AAY06526), of Drosophila non-receptor
XX CC serine/threonine kinase WART. hWART1 cDNA was isolated from a
XX CC human bone marrow cDNA library using PCR fragments of hWART1 as
XX CC probes. 3 Polymorphisms were identified in the hWART1 gene: (1) a

CC nucleotide 978, resulting in an Ala/Gly change; (2) at nucleotide
 CC 1840, silent; and (3) at nucleotides 3252-3253, comprising a
 CC deletion of 2 adenosines, resulting in a C-terminal truncation of
 CC hWART1 in the putative kinase domain. The latter frameshift
 CC mutation was observed in 2 independent clones from human bone
 CC marrow cDNA. Truncation of hWART1 could play a role in disease
 CC progression. hWART1 shows strong expression in cell lines from
 CC non-small cell lung cancer, ovarian tumours, central nervous system
 CC tumors, renal tumours and breast tumours, and may provide a target
 CC for oncology drug development. Nucleic acids encoding full-length
 CC hWART1 and hWART1 polypeptides lacking one or more of amino acid
 CC segments 12-45, 55-151, 236-377, 404-520, 555-559, 601-702, 691-998
 CC and 1011-1086, or lacking one or more of the N-terminal domain,
 CC catalytic domain, or C-terminal domain are claimed, as well as
 CC hWART2 sequences (AY87397), hWART1 and hWART2 polypeptides,
 CC antibodies, a method for identifying modulators of hWART function,
 CC and use of such modulator compounds to treat an abnormal condition
 CC involving hWART signal transduction, especially cancer. Probes for
 CC detection of hWART nucleic acids are also claimed.
 XX
 SQ Sequence 7382 BP; 2291 A; 1458 C; 1417 G; 2216 T; 0 other;

Query Match 77.2%; Score 2481.4; DB 20; Length 7382;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 2768; Conservative 0; Mismatches 386; Indels 15; Gaps 4;

QY 1 GTCAACATTCATACCAACAAAGCTGGAAGGTTCTAAAGAGTCTAGTTCCT 60
 DB GTGACGAATTCAGTTAACCCGAACAGAGCTGGAAGGTTCTAAAGAAATCCTTAGTTCCT 955
 QY 61 CAGAGACACGGCCCATCTCTAGGAGAAATGTGTGTTATCTGTTCTGAAAGCCCAACTCA 120
 DB 956 CAGAGGATGCGCCCATCTAGCAGAAAGTGTGCGCTATCATCTCAGAGTCCCAACTCA 1015
 QY 121 CAGCGGATGTAGAGACCTCTGTCTGGATCGCGCATTCAGCATTTGCTCAGCTCAC 180
 DB 1016 CAGACAGATGTAGAGACCTCTGTCTGGATCTGGTATATCAGCATTTGTTCAAGCTCAC 1075
 QY 181 CCAAGCAATGACAGAGAGTGAACCCGCCACCACTCAAGTTAGGAGTGTACTCCT 240
 DB 1076 CCTAGCAACGGCAGAGAGTGAACCCGCCACCACTCAAGTAAGGAGTGTACTCCT 1135
 QY 241 CCACCACTCCGAGAGCCACACCCACCTCCCGAGGACCACTCCCTCCCTCCCTCCCTCA 300
 DB 1136 CCACCACTCCGAGAGCCACACCTCCCGAGGAGTGAACCTCCACCTCCCTCCCTCA 1195
 QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGAGTACGTAATCTCCCGA 360
 DB 1196 TGGGAACCAACTCTCAACAAAGCGTATCTGGAAACATGGATACGTAATCTCCCGA 1255
 QY 361 ATCTCCCTGTTCACCTGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCACT 420
 DB 1256 ATCTCTCTGTCACCTGGGCAATGCAAGAGGGTATCTCCACCACTCTCAACACT 1315
 QY 421 TCTCCCATGATCCCTAGCCAGGCTCAGAGGCCATTAGTCTGTCTCAGTTGGTAGA 480
 DB 1316 TCCCCCATGATCTCTTAATCAAGGACAGAGAGGCAATAGTCTGTCTGTGTGGCAGA 1375
 QY 481 CAACCCATCATCTCAGAGTACTAGCAAAATTTAACTTTACACCGGGCGACCTGGAGTT 540
 DB 1376 CAACCAATCATCTCAGAGTCTTAGCAAAATTTAACTTTTCAACAGGAGACCTGGAGT 1435
 QY 541 CAGAAATGGTGTGCTGAGTCTGATTTTATCGTGACCAAAA ---TGTCCCACTGGTTCT 597
 DB 1436 CAGAAATGGTGTGACAAACTGATTTTCATGATACACAAAATGTTGTCCTGCTGGCACT 1495
 QY 598 GTGACTGGGAGCCACCACTCCATATCTCTGACCCAGCTTAATGGACAAAGCCCTCT 657
 DB 1496 GTGAATGGGAGCCACCACTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTCT 1555
 QY 658 GCTTTTACAAACAGGGGCTGTCTGCTCCCACTCATTCGCAATGGAAGGTTCTTCAG 717
 DB 1556 GCTTTTACAAACAGGGGCTGTCTGCTCTCTCATATACAAATGGAAGTATTCCTTCAG 1615

QY 718 TCGATGATGCTGCCCAACAGGAACAGCTATAACATGGAGCTTTATAATATATGTCCT 777
 DB 1616 TCTATGATGCTGCCCAACAGGAATAGTATACATGGAAGTATATACATTTAGTACCT 1675
 QY 778 GGAAGTCAACAGAGCTGGCCCGAGTGTCTTCTGCTCTGCGCAGTCTATCCCAAGCGGT 837
 DB 1676 GGAAGTCAACAAATTTGGCTCAGTCTATCTCTGCTCCAGCCAGTCTATCCCGAGCACT 1735
 QY 838 GGGCATGAATTCCTACATGGAACCTACATACAGTCTGAGTCAAAATCTTTTATAAC 897
 DB 1736 GGGCATGAATTCCTACATGGAACCTACATACAGTCTGAGTCAAAATCTTTTATAAC 1795
 QY 898 CCATTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTTACGCTTCTGCCACTACAGTCACT 957
 DB 1796 CCATTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTTACGCTTCTGCCACTACAGTCACT 1855
 QY 958 GGCATCACACCGCTCTTATTAACAGCCGCTGAAGAGATGCGCTCTGAAACAGAG 1017
 DB 1856 GCAATTTACACAGCTCTTATTAACAGCCGCTGGAAGAGTATGCGTGTATTAACAGAG 1915
 QY 1018 CTGCAGCTGCTTTAGCCCAACCCCATCTCTTGGATGCCACAGCAGCTTCAGACTGTT 1077
 DB 1916 CTACAGACTGCTTTAGCACCCTACACACCTCTTGGATACCACAGCAATTCAACTGTT 1975
 QY 1078 CAGCCTTACCCCTTTTCTGAGGTACAGTTCAGTGTGCTGTCTATCCACTGTTGCT 1137
 DB 1976 CAAACCCAGCTCTTTCTGAGGGAACCGCTTCAATGTGACTGTGATGCCACTGTTGCT 2035
 QY 1138 GAAGTCTCAAGCTATCAAGTCCACCCGCTTATCCAAACATCTGTACACCAAAAC 1197
 DB 2036 GAAGTCTCAAACTATCAAGGACCCACCCACCTACCCAAACATCTGTGACCAAAAC 2095
 QY 1198 CCATCTGCTCCCTCATATGAGTCAAGTAAAGTCAAGCTGCAAGAGTGAACAGCTAGCTTA 1257
 DB 2096 CCATCTGCTCCCTCATATGAGTCAAGTAAAGTCAAGCTGCAAGAGTGAACAGCTAGCTTA 2155
 QY 1258 CCAAGGAAGATGATAGTGAAGAGTGGGACAGTGGTGTCTCTGGGATTAAGAAAG 1317
 DB 2156 CCAAGGAAGATGAGAGTGAAGAGTGGGATGAGTGAAGTGGGATTAAGAAAG 2215
 QY 1318 AAACAGATTAACACTTCACTCTGTTGCGGAAACAAAGAAAGTGAAGAACGAAGA 1377
 DB 2216 AAACAGATTAACACTTCACTCTGTTAGTGAAGAAACAAAGAAAGTGAAGAACGAAGA 2275
 QY 1378 GAGTCTCGGATTCAGAGTCTACTCCCAAGCGCTTTAAGTCTTCTGAGAGCAGCTA 1437
 DB 2276 GAATCTCGTATTCAAAGTATTTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGA 2335
 QY 1438 GAGAACGCTCTGAGTCTATCAGCAGCGCTCTGATCGGAAGAGCAGCTAGAAAATGAA 1497
 DB 2336 GAAATGTACTCAATCTCATCAGCAGCGCTCTACATCGTAAACAAANTTAGAATGAA 2395
 QY 1498 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAATGAGAAAGTGTCTTCCAG 1557
 DB 2396 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAATGAGAAAGTGTCTTCCAG 2455
 QY 1558 AAAGAGTCTAACTATATCTCTTTAAAGGCTTAAATGACAAAGTCTATGTTTGAAG 1617
 DB 2456 AAAGAACTAATTAACATCCCTCTTTAAAGGCTTAAATGACAAAGTCTATGTTTGAAG 2515
 QY 1618 ATAAAGCATTAGAATAGGAGCGTTGGTGAAGTCTGTAGCAAGAAAGTCTGATACT 1677
 DB 2516 ATAAAGCATTAGAATAGGAGCGTTGGTGAAGTCTGTAGCAAGAAAGTCTGATACT 2575
 QY 1678 AAAGCTTTGTATGCACAAAGACTCTTTCGAAAGAAAGACGCTTCTGCTCCGAAATCAGGTG 1737
 DB 2576 AAGCTTTGTATGCACAAAGACTCTTTCGAAAGAAAGACGCTTCTTCTTCCGAAATCAGGTG 2635
 QY 1738 GCTCATGTGAAGGGAGAGGATATCTCTAGCAGAGCCGACCAATGAGTGGTGGTCCG 1797
 DB 2636 GCTCATGTGAAGGGAGAGGATATCTCTGCTGAAGCTGACAAATGAATGGTGGTGGTTCGT 2695

QY 1798 CTGTACTACTCTTCCAGCAAGACAACTGTGACTTTGTGATGGACTACATCTCTGGG 1857
 Db 2696 CTATATATTATTCATTCACCAAGATAAGACAAATATATCTTTGTAAGACTACATCTCTGGG 2755
 QY 1858 GGGGATATGATGAGCCTATTATAGATGGGCATCTTCTGAAATCTGGCAGATTC 1917
 Db 2756 GGTGATATGATGAGCCTATTATAGATGGGCATCTTCTGAAATCTGGCAGATTC 2815
 QY 1918 TACATAGCAGAACTTACCTGTGCAAGTTCGAAAGTGTTCATAAAATGGGTTTATTCATAGA 1977
 Db 2816 TACATAGCAGAACTTACCTGTGCAAGTTCGAAAGTGTTCATAAAATGGGTTTATTCATAGA 2875
 QY 1978 GATATTAACCTGATAACATTTTATGACCGGTGATGCCATATTAATAATGACTGACTTT 2037
 Db 2876 GATATTAACCTGATAATTTTATGATGATGATGATGATGATGATGATGATGATGATGAT 2935
 QY 2038 GGCTGTGCACTGCTTCAGATGGACATGACTCTCAAGTACTACAGAGTGGGATCAC 2097
 Db 2936 GGCTGTGCACTGCTTCAGATGGACATGACTCTCAAGTACTACAGAGTGGGATCAC 2995
 QY 2098 CCAGCGCAAGATAGCATGGATTCAGTAACGAATGGGAGATCTTCCAAATTTGCGGTGT 2157
 Db 2996 CCAGCGCAAGATAGCATGGATTCAGTAACGAATGGGAGATCTTCCAAATTTGCGGTGT 3055
 QY 2158 GGGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACCAGCGATGTAGCC 2217
 Db 3056 GGAGACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACCAGCGATGTAGCA 3115
 QY 2218 CATTTCTGGTGGACTCCCAATATATTCGACCTGAAGTGTCTACCTCGAAGAGGATAT 2277
 Db 3116 CATTTCTGGTGGACTCCCAATATATTCGACCTGAAGTGTCTACCTCGAAGAGGATAT 3175
 QY 2278 ACACAGCTGTGCTGAGTGGTGGTGTGTTTCTTCTGGAATCTTCTGTTGGGACAA 2337
 Db 3176 ACACAGTGTGCTGAGTGGTGGTGTGTTTCTTCTGGAATCTTCTGTTGGGACAA 3235
 QY 2338 CCTCTTTCTTGGCACAACCCCAATAGAAAACACAAATGAAGTATCATCTGGCAACT 2397
 Db 3236 CCTCTTTCTTGGCACAACCCCAATAGAAAACACAAATGAAGTATCATCACTGGCAACA 3295
 QY 2398 TCTTACACATCTCTCTCAAGCTAAGCTGAGTCTGAGCCTCTGACCTCATATATCAAA 2457
 Db 3296 TCTTCTTCAATTCACCAACAGCTAAACTCAAGTCTGAGCTCTGAGCTCTTATATTA 3355
 QY 2458 CTGTGTCGAGCAGACAGACCGCTCGGCAAGACGGTCTGATGAGATAAAGCTCAT 2517
 Db 3356 CTGTGTCGAGCAGACAGACCGCTCGGCAAGACGGTCTGATGAGATAAAGCTCAT 3415
 QY 2518 CCATTTTAAAGACCATCGATTTCTCTAGTGTCTGAGACAGCAGTCTGCTTCATACATC 2577
 Db 3416 CCATTTTAAAGACCATCGATTTCTCTAGTGTCTGAGACAGCAGTCTGCTTCATACATC 3475
 QY 2578 CCTAAATCAGCATTCACAGATACATCAATTTTGGACCTGTGATCTGATGATAATG 2637
 Db 3476 CCTAAATCAGCATTCACAGATACATCAATTTTGGACCTGTGATCTGATGATAATG 3535
 QY 2638 TGGAGCATGTCAGCAGGAGGAGAAATATCAGTACACTCTGAGCGGATGTATATAAAT 2697
 Db 3536 TGGAGCATGTCAGCAGGAGGAGAAATATCAGTACACTCTGATGATGATGATGATGAT 3595
 QY 2698 GGAAGCAGCCCGAGCAGCTTTCTATGAGTTCACCTTTCGGAGGTTTCTGATGACAA 2757
 Db 3596 GGAAGCAGCTCGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3655
 QY 2758 GGCTACCATATATTTATCCAAAGCCTATTCAGTATGAATACATTCATTCACAGGCTCA 2817
 Db 3656 GGCTACCATATATTTATCCGAAGCCTATTCAGTATGAATACATTCATTCACAGGCTCA 3715
 QY 2818 GAACAACAGTCTGATGAGATGATCAACACAGAGCTCCGATGGAAACACCCGAGATCTA 2877
 Db 3716 GAGCAGAGTCCGATGAGATGATCAACACAGAGCTCCGATGGAAACACCCGAGATCTA 3775
 QY 2878 GTGTATGTTTAAATAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 2933

Db 3776 GTATATGTTTAAACACACTAGTAAATAAATGTAATGAGGATTTGTAAGGCGCTGAAATG 3835
 QY 2934 CAGGGTTTTTGAAGTCTTTGAG---AAATATGCAAAATGTGACAGAG---TTTGTGT 2985
 Db 3836 CGAGTGTTTTGAGGTTCTGAGAGTAAATATGCAAAATATGACAGAGCTATATGTTGT 3895
 QY 2986 GCTCTGTCTACAAATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3045
 Db 3896 GCTCTGTCTACAAATATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3955
 QY 3046 ATTCACCCCTTTTAAATCAGTAATTTAGAAAAATTTGTTATAAGGAAAGTAAATATGAA 3105
 Db 3956 ATTCACCCCTTTTAAATCAGTAATTTAGAAAAATTTGTTATAAGGAAAGTAAATATGAA 4015
 QY 3106 CTGACTATATAGTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3154
 Db 4016 CTGAAATATATAGTCAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4064

RESULT 6

AA32981
 ID AAX32981 standard; DNA; 2442 BP.

XX AAX32981;

XX 30-JUN-1999 (first entry)

XX Human warts protein encoding DNA.

XX Human; warts protein; Drosophila; protein kinase; epithelial cell;
 KW cell regulation; cell proliferation; cell differentiation; tumour; ss.

XX Homo sapiens.

XX WO9915558-A1.

XX 01-APR-1999.

XX 24-AUG-1998; 98WO-JP03739.

XX 24-SEP-1997; 97JP-0258689.

XX (SUME) SUMITOMO ELECTRIC IND CO.

XX Kishimoto T; Negamine Y, Nishiyama Y, Niwa S, Saya H;

XX WPI; 1999-244389/20.

XX P-PSDB; AAY06890.

XX Gene of vertebrates homologous to a Drosophila gene

XX Claim 7; Page 44-52; 68pp; Japanese.

XX The invention relates to a human warts protein, homologous to Drosophila
 CC warts gene expression product. The warts expression product is a protein
 CC kinase involved in the regulation of proliferation and differentiation of
 CC epithelial cells and suppression of tumour formation. The gene, its
 CC expression product and antibodies are useful in the study of the
 CC mechanisms of tumour development. The present sequence represents a DNA
 CC encoding a human warts protein.

XX Sequence 2442 BP; 777 A; 528 C; 491 G; 646 T; 0 other;

Query Match 60.6%; Score 1946.8; DB 20; Length 2442;
 Best Local Similarity 87.9%; Pred. No. 0;
 Matches 2146; Conservative 0; Mismatches 292; Indels 4; Gaps 2;

QY 331 TCTGGACATGGAGTACGTATCTCCGAAATCTCCCTGTTCACCTGGGCGTGGCAG 390

Db 1 TCTGGAAACATGGAAATAGTATCTCCGAAATCTCTCTCTCCACCTGGGCGATGGCAA 60

QY 391 GAGGGGTACCTCCACCACTCTTACCACTTCTCCATGAATCCCTCCCTAGCCAGCTCAG 450

Fri Jan 17 11:16:58 2003

1141	CTACATCTTAAACAAATTTAGAGAAATGAATGATCGGGTTGGATTATCTCAAGATGCC	1200
1528	CAGGATCAAAATGAGAAGATGCTTTGCCGAGAAGAGTCTAACTATATATCGTCTTAAAGG	1587
1201	CAGGATCAAAATGAGAAGATGCTTTGCCGAGAAGAGTCTAACTATATATCGTCTTAAAGG	1260
1588	GCTAAATGACAAAGTCTATGTTTGAAGATAAAGACATTAAGGAATAGGAGCGTTGGT	1647
1261	GCTAAATGACAAAGTCTATGTTTGAAGATAAAGACATTAAGGAATAGGAGCGTTGGT	1320
1648	GAAGTCTCTAGCAGAAAGTGGATTAAGCTTTGTATGCAACAAAGACTCTTCCA	1707
1321	GAAGTCTCTAGCAGAAAGTGGATTAAGCTTTGTATGCAACAAAGACTCTTCCA	1380
1708	AAGAAAGACGCTCTGCTCCGAAATCAGTGGCTCATGTGAAAGCGGAGGATATCTTA	1767
1381	AAGAAAGACGCTCTGCTCCGAAATCAGTGGCTCATGTGAAAGCGGAGGATATCTTA	1440
1768	GCAGAAAGCGGAGTGGTGGTGGTCTGCTCTATATATATATATATATATATATATAT	1827
1441	GCTGAAGCTGACAAATGAATGGTGGTGGTCTGCTCTATATATATATATATATATAT	1500
1828	TGCTACTTTGCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1887
1501	TTATATCTTTGTAATGGAGTACATCTCTGGGGTGGTGGTGGTGGTGGTGGTGGTGGT	1560
1888	GCATCTTTCTGAAATCTGGCAGATTTACATAGCAGAACTTACCTGTGCAAGTTGAA	1947
1561	GCATCTTTCTGAAATCTGGCAGATTTACATAGCAGAACTTACCTGTGCAAGTTGAA	1620
1948	AGTGTTCATTAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2007
1621	AGTGTTCATTAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1680
2008	CGTGTGGCCATTAATAATTTGACTGACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	2067
1681	CGTGTGGCTCATATAATTTGACTGACTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	1740
2068	GACTCCAGTACTACAGAGTGGGATCACCCAGGAGATGAGATGAGATGAGATGAGATGAG	2127
1741	GATTCTAAGTACTATCAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1800
2128	GAATGGGAGATCTTCCAAATTTGCTGGTGGGAGACAGACTGAAGCCCTGAGCGGAGA	2187
1801	GAATGGGAGATCTTCCAAATTTGCTGGTGGGAGACAGACTGAAGCCCTGAGCGGAGA	1860
2188	GCTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2247
1861	GCTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1920
2248	GCACCTGAAGTGGTCTTACGAAAGGATATACACAGCTGTGCTGCTGCTGCTGCTGCTG	2307
1921	GCACCTGAAGTGGTCTTACGAAAGGATATACACAGCTGTGCTGCTGCTGCTGCTGCTG	1980
2308	GTTATCTTTGTAATGTTGGTGGGACACCTCTTCTTGGCAGCAAAACCCATTTAGAA	2367
1981	GTTATCTTTGTAATGTTGGTGGGACACCTCTTCTTGGCAGCAAAACCCATTTAGAA	2040
2368	ACACAAATGAAGTGGTCTTACGAAAGGATATACACAGCTGTGCTGCTGCTGCTGCTG	2427
2041	ACACAAATGAAGTGGTCTTACGAAAGGATATACACAGCTGTGCTGCTGCTGCTGCTG	2100
2428	AGTCTGAGGCTCTGACCTCATTTATCAACTGCTGAGAGCAGAGCAGAGCAGAGCAGAG	2487
2101	AGTCTGAGGCTCTGACCTCATTTATCAACTGCTGAGAGCAGAGCAGAGCAGAGCAGAG	2160
2488	AGAAGGCTGCTGATGAGATTAAGGCTCATCTTATTTTAAAGCCATCTGCTTCTAGT	2547
2161	AGAAGGCTGCTGATGAGATTAAGGCTCATCTTATTTTAAAGCCATCTGCTTCTAGT	2220
2548	GATCTGAGACAGCTGCTGCTTATCATCTTAAATCAGCATCCAAACAGATACATCC	2607
2221	GACCTGAGACAGCTGCTGCTTATCATCTTAAATCAGCATCCAAACAGATACATCA	2280
61	GAGGGCTCTCCACCACTTTTCAACAGTCTCCCCCAATGAATCCTCTTAATCAAGGACAG	120
451	AGGGCCATAGTCTGCTTCCAGTGGTAGACAACCCATCATCATGAGAGTACTAGCAAA	510
121	AGAGGCAATAGTCTGCTTCCAGTGGTAGACAACCCATCATCATGAGAGTACTAGCAAA	180
511	TTTAACTTTACACAGGGGACCTGAGTGTAGATGGTGGTGGTGGTGGTGGTGGTGGTGGT	570
181	TTTAACTTTACACAGGGGACCTGAGTGTAGATGGTGGTGGTGGTGGTGGTGGTGGTGGT	240
571	GTGACCAAAA---TGTCCTCACTGCTGTGACGTCGGCAGCAGCAGCAGCAGCAGCAGCAG	627
241	ATACACCAAAATGTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	300
628	CTGACCCAGCTAATGACAAAGCCCTCTGCTTTCAACAGGGGCTTCTGCTGCTCA	687
301	CTGACAGCAGTAAATGACAAAGCCCTCTGCTTTTCAACAGGGGCTTCTGCTGCTCA	360
688	CCATCTGCGCATGAAAGCTTCTCTCAGTCTGATGATGGTGGTGGTGGTGGTGGTGGTGGT	747
361	TCGTATATCAAAAGGAGTATTCCTCAGTGTATGATGGTGGTGGTGGTGGTGGTGGTGGT	420
748	AACATGAGCTTTTATAATATTAATGTCCTGGAGTCAACAGCTGGCCGCTGCTGCT	807
421	AACATGAGCTTTTATAATATTAATGTCCTGGAGTCAACAGCTGGCCGCTGCTGCTGCT	480
808	TCTGCTCTGCGCAGTATCCCAAGCGGTGGGATGAAATTTCTACATGCGCAACCTAAC	867
481	TCTGCTCTGCGCAGTATCCCAAGCGGTGGGATGAAATTTCTACATGCGCAACCTAAC	540
868	ATACAGTGGTCAAAATCTTTTAAATACCCATTAGAAGTAGAGCAAGTCACTGCT	927
541	ATACAGTGGTCAAAATCTTTTAAATACCCATTAGAAGTAGAGCAAGTCACTGCT	600
928	AAATCTCAGCTTCTGCGCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	987
601	AAATCTCAGCTTCTGCGCAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	660
988	GTGAAAGCATCGGCTCTGAAACAGAGTGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1047
661	GTGAAAGTATCGCTATTAACCAAGAGTGGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	720
1048	TCTTGGATCCAGCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	1107
721	TCTTGGATCCAGCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	780
1108	TCAAGTGTGCTGTATCCACCTGTTGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	1167
781	TCAAGTGTGCTGTATCCACCTGTTGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	840
1168	CCCTTATCCAAACATCTGCTACACCAAAACCCATCTGCTCCCTCCATATGAGTCAAGT	1227
841	CCCTTATCCAAACATCTGCTACACCAAAACCCATCTGCTCCCTCCATATGAGTCAAGT	900
1228	AAGCCCTGCAAGATGAACAGCTAGCTTACCAAGAGAGATGATAGTGAAGAGTGG	1287
901	AAGCCCTGCAAGATGAACAGCTAGCTTACCAAGAGAGATGATAGTGAAGAGTGG	960
1288	GACAGTGGTACTCTGGGATTAAGAAAGAAAGAGATTAACACTTCACTATCACTGTT	1347
961	GAAATGTTGATGTTGGGATTAAGAAAGAAAGAGATTAACACTTCACTATCACTGTT	1020
1348	CGGAAACAAAGATGAAGAACGAGAGAGTCTCGGATTCAGATTCAGATTCCTCCACAG	1407
1021	AGGAAACAAAGATGAAGAACGAGAGAGTCTCGGATTCAGATTCAGATTCCTCCACAG	1080
1408	GCCTTTAAGTCTTCAATGAGCAGCAGTATGAGAGACCTCTGAGTCTCATCAGCAGCT	1467
1081	GCATTTAAATCTTTATGAGCAGCAGTATGAGAGACCTCTGAGTCTCATCAGCAGCT	1140
1468	CTGCACTCGAAGCAGCTAGAAATGAAATGATGGGGTGGATTTATCTCAAGATGCC	1527

CC involved in oncogenesis, cancer cell proliferation or metastasis.
CC Depending on their biological activities, polypeptides and nucleotides of
CC the invention are useful for preventing, treating or ameliorating medical
CC conditions, e.g., by protein or gene therapy. Such conditions include
CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC proliferative retinopathy, atherosclerosis, coronary heart disease,
CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC vascular growth. Polypeptides involved with tissue regeneration and
CC repair (or nucleic acids encoding them) may be used to promote wound
CC healing (e.g., of burns, incisions and ulcers), while those with
CC immunomodulatory activities may be used in the treatment of viral,
CC bacterial and fungal infections in addition to immune disorders.
CC Polypeptides with growth factor activity may be used in cell cultures to
CC promote cell growth. For example, such polypeptides may be used to
CC manipulate stem cells in culture to give rise to neuroepithelial cells
CC that can be used to augment or replace cells damaged by illness,
CC autoimmune disease or accidental damage. The polypeptides and nucleotides
CC may also be used in the diagnosis of the above conditions, and in drug
CC screening techniques. The present sequence represents a cDNA encoding a
CC novel human polypeptide of the invention.
XX Sequence 1357 BP; 436 A; 254 C; 285 G; 382 T; 0 other;
SQ

Query Match 31.5%; Score 1012; DB 22; Length 1357;
Best Local Similarity 85.7%; Pred. No. 3.8e-242;
Matches 1164; Conservative 0; Mismatches 180; Indels 14; Gaps 3;

QY 1542 AAGATGCTTTGCCAAGAGCTCACTATATTCCTCTTAAAGGGCTTAAATGGACAA 1601
DB 1 AAGATGCTTTGGCCA-AAGATCTTAACTACCTCGCTTAAAGGGCTTAAATGGACAA 59
QY 1602 GTCTATGTTTGAAGATAAGACATTAGAATAGGAGCGTTTGGTGAAGTCGTCTAGC 1661
DB 60 GTCTATGTTTGAAGATAAGACATTAGAATAGGAGCGTTTGGTGAAGTCGTCTAGC 119
QY 1662 AAGAAAGTCGATCTAAAGCTTTGTATGCAACAAGACCTCTTCAAGAAAGACGTTCT 1721
DB 120 AAGAAAGTAGATACTAAGGCTTTGTATGCAACAAGACCTCTTCAAGAAAGACGTTCT 179
QY 1722 GCTCCGAAATCAGTGGCTCATGTGAAGCGGAGAGGATATCTAGCAGAACCCGACAA 1781
DB 180 TCTTCGAATCAAGTCGCTCATGTGAAGCGTGAAGAGATATCTGGCTGAAGCTGACAA 239
QY 1782 TGAGTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1841
DB 240 TGAATGGGTAGTTCGCTCTATATTTATTCCTCAAGATAGGACAAATTTATACCTTTGAAT 299
QY 1842 GGACTACATCTCCGCGGGGATATGATGAGCCTATTATTAATAGATGGCATCTTCTCTGA 1901
DB 300 GGACTACATCTCCGCGGGGATATGATGAGCCTATTATTAATAGATGGCATCTTCTCTGA 359
QY 1902 AAATCTGGCAGATTCTACATAGCAGAACTTACCTGTCAGTGGAAAGTGTTCATAAAT 1961
DB 360 AAGCTGGCAGATTCTACATAGCAGAACTTACCTGTCAGTGGAAAGTGTTCATAAAT 419
QY 1962 GGGTCTTATTCATAGAGATATTAAACCTGTAAACATTTTGAATGACCGTGTATGCCATAT 2021
DB 420 GGGTCTTATTCATAGAGATATTAAACCTGTAAACATTTTGAATGACCGTGTATGCCATAT 479
QY 2022 TAAATGACTGACTTTGGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2081
DB 480 TAAATGACTGACTTTGGCTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 539
QY 2082 CCAGAGTGGGATCACCCACCGGAGATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2141
DB 540 TCAGAGTGGTGGATTCACCGGAGATAGATGATGATGATGATGATGATGATGATGATGATGATGAT 599
QY 2142 TTCCAAATGTCGGTGGGAGAGACTCAAGCCTGAGCGGAGAGTGTCTGCCAGCA 2201
DB 600 CTCAGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 659
QY 2202 CCAGCGATGTCTAGCCCAATTCCTGCTGGTGGGAGTCCCAATTTATTTGACCTGAGTGTCT 2261

QY 2656 GAGGAAATATCAGTACACTCT-GAGCGATGGTATAAATGGAAGCACCOCGACGA 2714
DB 1261 GAGGAAATCTAATGACACTCTCAATGGATGGTATAAATGGAAGCACCOCGACGA 1320
QY 2715 CGCTTCTATGAGTTCACCTTCGGAGGTTTGTGATGACAATGGCTACCCATA 2768
DB 1321 TGCATTCATGAATTTACCTTCGGAAGGTTTGTGATGACAATGGCTACCCATA 1374

RESULT 8
ABA08740
ID ABA08740 standard; cDNA; 1357 BP.
AC ABA08740;
DT 11-JAN-2002 (first entry)
XX Human large tumour suppressor 1 homologue cDNA, SEQ ID NO:516.
DE Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritis; haemostatic; antiarteriosclerotic;
KW cystostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnery; antitumor; ss.
XX Homo sapiens.
XX WO200157188-A2.
PN 09-AUG-2001.
PD 05-FEB-2001; 2001WO-US03800.
PF 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
XX (HYSE-) HYSEQ INC.
PA Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-457740/49.
XX P-PSDB; ABB11496.
XX Human proteins and DNA encoding sequences useful for preventing,
PT treating or ameliorating a medical condition in a mammalian subject
PT e.g. arthritis and cancer
PS Claim 1; Page 554-555; 1963pp; English.
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC sequences ABB08225-ABB09574 represent nucleic acids encoding them. The
CC invention also relates to vectors and recombinant host cells comprising a
CC nucleotide of the invention, methods of producing the novel polypeptides,
CC antibodies against the polypeptides, methods of detecting the nucleotides
CC or polypeptides in a sample, and methods of identifying compounds which
CC bind to polypeptides of the invention. Although novel, many of the
CC polypeptides of the invention have homology to known proteins, thereby
CC giving an insight into their probable biological activities, and hence
CC potential therapeutic applications. The polypeptides of the invention may
CC have various activities, including cytokine, cell proliferation or cell
CC differentiation activities; stem cell growth factor activity;
CC haematopoiesis regulatory activity; tissue growth activity;
CC immunomodulatory activity; activin- or inhibin-related activities;
CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC thrombolytic activities; receptor or ligand activities; or may be

```

Db 660 CCAGCGATGCTAGACATCTCTTGGTGGGACTCCCAATATATATGCACTGAAAGTGT 719
Qy 2262 ACTGCGAAGAGATATACACAGCTGTGACCTGGTGGAGTGTGTTATCTTTGTGA 2321
Db 720 GCTAGCAAGAGATACACAGTGTGTGATTTGGTGGAGTGTGTTATCTTTGTGA 779
Qy 2322 AATGTTGGTGGACAACTCTCTTCTTGGCAAAACCCCATTTAGAAACACAAATGAAGT 2381
Db 780 AATGTTGGTGGACAACTCTCTTCTTGGCAAAACCCCATTTAGAAACACAAATGAAGT 839
Qy 2382 TATCATCTGGCAAACTCTTACACATCCCTCAAGCTAGCTAGCTGCTGAAGCTC 2441
Db 840 TATCAACTGGCAAACTCTTACATCTCCACCAAGCTAACTCACTCTCTGAAGCTC 899
Qy 2442 TGACCTCATTTATCAAACTGTGTCGAGGACCAAGACCCCTCGGCAAGAGGGTCTCA 2501
Db 900 TGATCTTATTAACATTTGCCGAGGACCCGGAAGTCTGTTAGGCAAGATGGTCTCA 959
Qy 2502 TGAGATAAGGCTCATCCATTTTAAAG-----CCATCGATTTCTCTAGTGA 2549
Db 960 TGAATAAAAGCTCATCCATTTTAAAGCAATTTGACTTCTCCAGTGACCTGAAGA 1019
Qy 2550 TGTGAGACAGAGCTGTCTTATACATCCCTTAAATACAGCATCCACAGATACATCAA 2609
Db 1020 CAGCAGGTGGCTTTCAAACTTCCCTTAAATACACACCCACCAAGATACATCAA 1079
Qy 2610 TTTCGA-CCCTGTTGATCTGATATAATTTGAGCGATGGCGAGGAGGAAATATCA 2668
Db 1080 TTTTGTGCTGTTGATCTGATATAATTTGAGCGATGGCGAGGAGGAAATATCA 1139
Qy 2669 GTGACACTCTGAGCGGATGGTATAAAATTTGGAAGCAGCCCGGAGGAGGAGGAA 2728
Db 1140 ATGACACTCTCAATGATGATATAAAATTTGGAAGCAGCCCGGAGGAGGAGGAA 1199
Qy 2729 TCACCTTTGGAGGTTTTTGTGATGACATGCTACCCATATAATTTATCAAGCTATTG 2788
Db 1200 TTACCTTCCGAGGTTTTTGTGATGACATGCTACCCATATAATTTATCAAGCTATTG 1259
Qy 2789 AGTATGATATATTCATTCACAGGCTCAGAACACAGCTGTGATGATGATGTT 2886
Db 1320 CAGGCTCAGAGATTAAATAATCGCATCTAGTATATGTT 1357

```

RESULT 9

AA42120
ID AA42120 standard; cDNA; 3155 BP.

AC AA42120;

XX 31-JAN-1997 (first entry)

XX M-lats2 gene encoding large tumour suppressor.

DE Mouse; m-lats2 gene; large tumour suppressor; fetal brain;

XX protein-serine/threonine-kinase; cell proliferation; antisense;

KW dominant-negative; cancer; degenerative disorder; trauma;

KW growth deficiency; therapy; antitumour; vulnary; diagnostic;

KW transgenic plant; transgenic animal; growth; senescence; ds.

XX Mus musculus.

OS Key

XX Location/Qualifiers

FT 1..2943

FT /*tag=

FT /product= m-lats2 protein

XX W09630402-A1.

PN

XX

03-OCT-1996.
26-MAR-1996; 96WO-US04101.
27-MAR-1995; 95US-0411111.
(UYUA) UNIV YALE.
Tao W, Wang W, Xu T, Yu W, Zhang S;
WPI; 1996-455275/45.
P-PSDB; AAW06287.
New isolated large tumour suppressor gene - used to develop prods.
for inhibiting cell proliferation or for enhancing proliferation.
Claim 8; Page 133-137; 215pp; English.
This sequence encodes a mouse large tumour suppressor m-lats2
protein, and has been isolated from a newborn mouse brain phage
lambda-ZAP cDNA library using a 2.2-kb DNA probe from the Drosophila
lats gene (AAT42117). A homologous mouse sequence has also been
isolated (m-lats, AAT42119). The gene encodes a putative protein-
serine/threonine-kinase, and inhibits cell proliferation and plays a
crucial role throughout development. Activators or inhibitors of
lats function (e.g. an antisense oligonucleotide or dominant-negative
lats fragment) may be used in therapy of cancer or other
proliferative disorders, degenerative disorders, trauma, growth
deficiency, etc., and fragments of the gene may be used as diagnostic
probes. A lats-inhibitor sequence may be expressed in a transgenic
plant or farm animal to confer increased growth and inhibit
senescence.
Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match 25.3%; Score 812.6; DB 17; Length 3155;
Best Local Similarity 65.1%; Pred. No. 3.3e-192;
Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;
Qy 919 CACTCTGCTTAATTCACGCTTCTGCCACTACAGTCACTGCCATCACACCCCTCTATT 978
Db 985 CAACCTGAGCCCTCACTGCGCGCCCAACACGGTGTTCCTTCCGCGCCGACACATC 1044
Qy 979 CAACAGCCCGTGAAGATGCGCTCTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCA 1038
Db 1045 CTTACACCTGTGAAGAGCTGCTGTGCGGCCGAGCCGACACAGCCGTGGGCCC 1104
Qy 1039 ACCCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTTTCTGAG 1098
Db 1105 TCGCACCCCGCTGGTGGCTGCGCCACAGCACCTGCCACTGAGAGGCTGGAGACGAAG 1164
Qy 1099 GTACAGCTTCAAGTGTGCTTCCACCTCTGTTGCTGAAGCTTCAAGCTATCAAGT 1158
Db 1165 GAGGCGAGCGAGGCCACACCCGCTGGATGTGGACTATGGCGGCTCCGAGCGAGGTGC 1224
Qy 1159 CCACCAACCGCTTATCCAAACATCTGTACACCAAAACCCATCTGCTCCCTCATATGAG 1218
Db 1225 CCACCGCTCTCGTATCCAAAGCACTTGTGCTGCCAGTAACTGAGCTGAGCAGCTACAGCTG 1284
Qy 1219 TCAGTAAGTAAGCCCTGCAAGATGAAGACCTAGCTTACCCCAAGAAAGATGATAGTGA 1278
Db 1285 GACCTGGACAGCCTGTGCACCAAGTGTGCAGCAGAGTCTGCGAGGGGCGACTGATAGAC 1344
Qy 1279 AAGAGTGGGACAGT-----GGTACTCTGGGATTAAGAAAGAAAGACAG 1323
Db 1345 GGGAGTGAACAGAGCCCAAAAGGTGCGAAGGAGGAGCAAGCTGGCAGAGACAAAAGCAG 1404
Qy 1324 ATTACAACCTCACCTATCACTGTTCCGAAACAAAGAGATGAAGAGAGAGAGTCT 1383
Db 1405 ATTACAGCTCCCGCTGCTGTCGCAAGATATACAGAGATGAAGAGAGAGAGTCT 1464
Qy 1384 CGGATTCAGAGTTACTCCCAAGGCTTTAAGTTCTTCTATGGAGCAGCAGCAGTAGAAC 1443
Db 1443 CGGATTCAGAGTTACTCCCAAGGCTTTAAGTTCTTCTATGGAGCAGCAGCAGTAGAAC 1443

Fri Jan 17 11:16:58 2003

Db 1465 CGATCAAGAGTTACTCCCTTTAGCCTTCAAAATCTTTCATGAGCAACACAGTGGAGAT 1524
Qy 1444 GTCTGAAGTCTCATCAGCAGCGTCTGCATCGGAAGACAGCTAGAAAATGAATGATG 1503
Db 1525 GTATCAAACTTACCAGCAGAGGTCAGCGGAGGCTACAGCTGAGCAGGAAATGGCC 1584
Qy 1504 CGGGTGGATATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTGCAGCAAGAG 1563
Db 1585 AAAGCTGGCTCTGTGAGCGCAGCAGGACAGATGAGGAAGATCTCTACCAAGAGGAG 1644
Qy 1564 TCTAACTATATTCGCTTTAAAGGGCTTAAATGGACAAGTCTATGTTTGTAAAGATAAAG 1623
Db 1645 TCTAACTACAAACCGCTTGAAGGGCCAGATGGACAAGTCCATGTTGTGAATAACAAG 1704
Qy 1624 ACATTAGGAATAGAGAGGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATTAAGCT 1683
Db 1705 ACTTAGCATCGTCTTGGGGAAGTGTCTGCTGCTGTAAGCTGGACACTACGCT 1764
Qy 1684 TTGTATGACAAAGACTCTTCGAAGAAAGAGTCTGCTCCGAATCAGGTGGCTCAT 1743
Db 1765 CTGTACGCCATGAAGACTCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGGCCAT 1824
Qy 1744 GTGAAGCGGAGAGGATATCTTAGCAAGAGCCGACAAATGAGTGGTGGTCCGCTGTAC 1803
Db 1825 GTCAAGCTGAGAGGACATCTGCTGAGCAGACAAATGAGTGGTGGTCAAACTCTAC 1884
Qy 1804 TACTCTTTCCAGGACAAAGGACAATCTGTGATGAGTACATCTCTGGGGGGAT 1863
Db 1885 TACTCTTTCCAGGACAAAGGACAGCTGTCTGATGAGTACATACACAGCGGGGAT 1944
Qy 1864 ATGATGACCTTAAATTAGAATGGCATCTTCTCGAAATCTGGCAGCATCTACATA 1923
Db 1945 ATGATGACCTTGTGATCAGATGAGGTCTCTCGACACCTGGCCGCTCTACAT 2004
Qy 1924 GCAGAACTTACCTGTGCAAGTGTGAAGTGTTCATTAATGGTGTTCATAGAGATAT 1983
Db 2005 GCAGAGTTGACCTGGCCATGAAAGTGTCCAGAGATGGCTTTATCCACCGGACATC 2064
Qy 1984 AAACCTGATACATTTTCATGACCGTATGCGCATATTAATGACTGACTTGGGTG 2043
Db 2065 AAGCTGACACATCTCATGACCTGATGATGATATTAAGCTGACAGATTTTGGCTC 2124
Qy 2044 TGCACTGGCTTCAGATGACACATGACTCCAGTACTACAGAGTGGGATCACCACGG 2103
Db 2125 TGCACTGGATTCAGGTGACTCACAATTCAGTACTACCAAGAGGGAACACATGAGA 2184
Qy 2104 CAAGATACATGATTTTCAGTAACCAATGGGAGATCTTCCAAATGTCGTGGGGAC 2163
Db 2185 CAGGACAGCATGGACCGGTGACCTCTGGGACGATGTTTCCAACTGCTGTGGAGAC 2244
Qy 2164 AGACTGAAGCCACTGGAGGAGAGTCTGCTGCCAGCAGCGATGCTAGCCCATCT 2223
Db 2245 AGTTAAAGACCTGGAGCAGAGGCGGAGAGCAGCAGCAGAGTGGCTGGCAGATCT 2304
Qy 2224 CTGGTGGGACTCCCAATTAATTCGACTGAAGTCTACTCGGACAGGATATACAG 2283
Db 2305 CTGTGGGACACCAATTTACATCTCGGAGGTCTTCCGCAAGGGTACACGAG 2364
Qy 2284 CTGTGTGACTGTGGAGTCTGTGTGTTTCTGAAATGTTGGTGGGACAACTCTCT 2343
Db 2365 CTCTGTGACTGTGGAGCGTGGTGTGATCTCTTTGAGATGCTGTGGCAGCGCT 2424
Qy 2344 TTCTTGGCACAACCCCATTTAGAAACACAAATGAAGGTTATCATCTGCAACCTCTTA 2403
Db 2425 TTCTTGGCCCCCACCACCCACAGACGACGACGAGTGAAGGTGATCAACTGGGAGACGCTG 2484
Qy 2404 CACATCCCTCTCAAGCTAAGCTGAGTCTGAGGCTCTGAGGCTCATATCAAACTGTGT 2463
Db 2485 CATATCCCTACGAGTGGAGCTTACGCGCTGAGGCCGAGACCTCATACAGAGCTGTGC 2544
Qy 2464 CGAGGACCAAGACACCGCTTCGCGCAAGAACAGGTGCTGTGATGAGATAAGGCTCATCCATTT 2523
Db 2545 TGCCTGGCTGACTGCCCTGGCAGGGATGGGCGAGATGAGCTCAAGGCACACCGCTTC 2604

Qy 2524 TTTAAGACCATCGAATTTCTCTAGTGTATCTGAGACAGCAGTCTGCTTATCATCATCCCTAAA 2583
Db 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGCAGGCTGCACCTAGCTCCCAACC 2664
Qy 2584 ATCAGCCATCCAAACAGATACATCCAAATTTCCGACCTGTGTGATCTGATAAATTTGGAGC 2643
Db 2665 ATCAGCCACCCCATGGACACCTCCAAATTTTGCAGCGTGGATGAAGAAAGCCCTGGCAC 2724
Qy 2644 GATGCGCAGGAGGAGAAATATCATCTGACACTCTGAGCGGATGATATAAAATGGGAAG 2703
Db 2725 GAGGCCAGCGGAGAG---AGGCCAAGCGCTGGGACACGCTGGCTCCGCCAGCAGCAAG 2781
Qy 2704 CACCCCGAGCAGCGCTTCTATGATTCACCTTTCGAGGTTTTCGAGTGAACAATGGCTAC 2763
Db 2782 CATCCAGACGCGCTCTATGATTCACCTTCCGCAAGTCTTCGATGACAAACGCTAT 2841
Qy 2764 CCATATAATTAATCCAAAGCCTATTGATGATGATATCATATT 2802
Db 2842 CCCTTCCGCTGCCGAAGCCCTCAGAGCCCGCAGAGAT 2880

RESULT 10

AAZ51507
ID AAZ51507 standard; DNA; 3155 BP

XX AAZ51507;

XX 21-JUN-2000 (first entry)

XX Mouse Lats2 (large tumour suppressor) DNA.

Mouse; Lats2; large tumour suppressor; cytostatic; vulnary;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..2943

XX /tag= a "Lats2 protein"

XX /product= 970..1920

XX misc_feature

XX /tag= b

XX /note= "This region is erroneously repeated in the

XX mouse Lats2 DNA sequence shown in figure 14"

XX misc_feature

XX /tag= c

XX /note= "This region is missing in the mouse Lats2

XX DNA sequence shown in figure 14"

XX WO200010602-A1.

XX PN

XX XX

XX PD

XX PF

XX 18-AUG-1999; 99WO-USI9068.

XX XX

XX 18-AUG-1998; 98US-0096996.

XX PR

XX 18-AUG-1998; 98US-0096997.

XX XX

XX (UYUA) UNIV YALE.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

XX Turenchalk GS, Stewart RA;

XX WPI: 2000-246496/21.

XX P-PSDB; AAY70392.

XX Use of lats: proteins, complexes of lats and cdc2 for treating cancer

PT

that is refractory to treatment by standard chemotherapy and radiation therapy, and disorders associated with aberrant levels of cdc2 activity

Claim 44; Pages 112-117; 134pp; English.

The present sequence is a DNA encoding mouse Lats2 (large tumour suppressor) protein which is a cell overproliferation inhibitor and a negative regulator of cell cycle-dependent kinase cdc2/cyclin A. The present sequence is useful for treating cancer that is refractory to standard chemotherapy or radiation therapy such as hyperplasia, metaplasia, or dysplasia, and disorders associated with aberrant levels of cdc2 activity. Conditions treated by promoting cdc2 function include degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds. An animal model preferably a mouse, in which a lats gene has been disrupted by homologous recombination, e.g. a lats knock-out mouse, is used for screening compounds that can be used to treat or prevent cancer, particularly skin cancer, soft tissue sarcomas and ovarian tumours, and disorders associated with pituitary dysfunction e.g. luteinizing hormone (LH) hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

Sequence 3155' BP; 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match : 25.3%; Score 812.6; DB 21; Length 3155;
Best Local Similarity 65.1%; Pred. No. 3.3e-192;
Matches 1237; Conservative 0; Mismatches 644; Indels 18;

2:

QY	919	C	A	C	T	G	C	T	A	A	T	T	C	T	C	A	G	C	T	T	C	G	C	A	T	C	A	C	C	G	C	T	C	T	A	T	T	978												
Db	985	C	A	A	C	T	G	A	C	C	C	T	C	A	T	T	G	C	C	C	C	C	A	A	C	G	G	T	C	A	C	C	G	C	G	C	A	C	A	T	C	1044								
QY	979	C	A	A	C	C	C	G	T	G	A	A	A	G	A	T	G	C	G	C	T	C	T	G	A	A	A	C	A	G	A	G	T	C	G	A	G	T	T	A	G	C	C	C	A	1038				
Db	1045	C	T	T	C	A	C	C	C	T	G	T	G	A	G	C	T	G	G	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	C	1104				
QY	1039	A	C	C	A	T	C	T	T	T	G	G	A	T	G	C	A	C	C	A	G	T	T	C	A	G	T	T	C	A	C	C	C	T	A	C	C	C	T	T	T	T	T	T	T	C	G	A	1098	
Db	1105	T	C	G	C	A	C	C	C	C	G	C	T	G	G	T	G	C	T	G	C	C	C	A	C	A	G	C	A	C	C	T	G	C	C	A	C	T	G	A	G	A	C	C	A	A	G	1164		
QY	1099	G	G	T	A	C	A	G	C	T	T	C	A	G	T	C	C	C	A	C	C	T	G	T	G	T	G	A	G	C	T	C	C	A	A	G	C	T	A	C	A	A	G	T	A	G	1158			
Db	1165	G	A	G	G	C	A	G	C	C	C	C	A	C	C	C	T	G	A	T	G	T	G	A	C	T	A	T	G	C	G	C	G	T	C	G	A	G	C	A	G	T	G	C	1224					
QY	1159	C	A	C	A	C	C	C	C	T	T	A	T	C	A	A	A	C	A	T	C	T	G	C	T	A	C	C	A	A	A	A	C	C	A	A	A	C	C	A	A	A	C	A	A	A	C	1218		
Db	1225	C	A	C	C	C	T	C	C	T	A	T	C	A	A	A	G	C	A	C	T	G	C	T	G	C	C	A	G	T	A	A	G	T	C	T	G	A	C	A	G	C	A	G	C	A	G	C	1284	
QY	1219	T	C	A	G	T	A	A	G	C	C	T	G	C	A	A	A	G	A	T	C	A	A	G	C	T	A	C	C	A	A	G	A	A	G	A	T	G	A	T	A	G	T	A	G	1278				
Db	1285	G	A	C	T	G	G	A	C	A	G	C	T	G	T	G	C	A	C	A	G	T	G	C	G	A	G	A	G	T	C	G	A	G	G	G	G	C	A	C	T	G	A	C	T	A	G	C	1344	
QY	1279	A	A	G	A	T	G	C	G	C	A	G	T	-----	-	G	T	G	A	C	T	G	G	G	A	T	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	A	1323
Db	1345	G	G	A	G	T	G	C	A	A	G	C	C	A	A	A	G	T	C	G	A	G	G	A	C	A	A	A	G	C	T	G	C	A	G	A	C	A	A	A	A	A	A	A	A	A	A	C	A	1404
QY	1324	A	T	T	C	A	A	C	T	T	C	A	C	T	T	C	G	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	T	1383
Db	1405	A	T	T	C	A	A	C	T	C	C	G	T	G	C	T	C	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	C	T	1464
QY	1384	C	G	A	T</																																													

D	b	1645	TCTAACTACAAACCGGTGAAGAGGGCCAAGATGGCAAGTCCATGTTTGTAAGAAATCAAG	1704
O	y	1624	ACATTAGGAATAGAGCGTTTGGTAGAGTCCTGTAGCAAGAAGATGCATACTAAAAGCT	1683
D	b	1705	ACTCTAGGCATCGTGCCCTTTGGGAAGTGTGCCCTGCTGTGAAGCTGGACACTCACGCT	1764
O	y	1684	TTGTATGCACAAGACTCTTCGAAAGAAAGACGTTCTCTCCGAAATCAGGTGGCTCAT	1743
D	b	1765	CTGTACGCCATGAAGACTCTCAGGAAGAAGGATGCTTGAACCGGAATCAAGTGGCCCAT	1824
O	y	1744	GTAAGCCGAGAGGATATCTTAGCAAGCCGACATAGTAGTGGTGGTCCGGCTGTAC	1803
D	b	1825	GTCAAGGCTCAGAGGACATCTCGCTGAAGCAGACATGAGTGGGTGGTCAAACCTCTAC	1884
O	y	1804	TACTCTTCCAGCACAAAGAACACTTGTACTTTGTGTGAGGACTACATTCTGTGGGGGAT	1863
D	b	1985	TACTCTTCCAGCAAGACACCCTGTACTTTGTGTGAGGACTACTATPACAGAGCGGGAT	1944
O	y	1864	ATGATGAGCCTATAATTAGAATTGGGCATCTTTCTCGTAAAAATCTGGCAGATTCTACATA	1923
D	b	1945	ATGATGAGCCTGTGTATCAGGATGGAGGTCTTCCCTGAGCACCTGGCCCGCTTCTACATT	2004
O	y	1924	GCAGAATTACCTGTGCGAGTTGAAAGTGTTCATAAAATGGGGTTTTATTTCATAGAGATATT	1983
D	b	2005	GCAGAGTTGACCTCGCCATTGAAAGTGTCCAAGATGGGCTTTATCCACGGGACATC	2064
O	y	1984	AACCTCGATAAACATTTTGATTGACCGTGATGCCATATTAATTCAGTACTTTGGCTTG	2043
D	b	2065	ANGCTTGACAACTACTCATCGACCTGGATGTGCATATTAAGCTGACAGATTTTTGGGCTC	2124
O	y	2044	TGCACCTGGCTTCAGATGGACATGACTCCAAAGTACTACCAGAGTGGGGATCACCCACGG	2103
D	b	2125	TGCACGTGATTGAGTGGACTCACAATTCCAAGTACTACCAGAAAGGACCAATCAGA	2184
O	y	2104	CAGATAGCATGGATTTTCAGTAACGAATGGGAGATPCCTTCCAATPTGCGTGTGGGAC	2163
D	b	2185	CAGCACAGCATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAACCTGTCGTGTGGAGAC	2244
O	y	2164	AGACTGAAGCCACTGGAGCGGAGAGCTGCTCCCGACGACACGAGGATGCTAGGCCANTTCT	2223
D	b	2245	AGGTTAAAGACCCTGGAGCAGAGGGCGCAGAAGCAGCACAGAGGTGCTTGGCACATTCT	2304
O	y	2224	CTGGTTGGGACTCCCATAATTATATTGCACCTGAAGTGCCTCCGAACAGGATATACACAG	2283
D	b	2305	CTTGTGGGACACCAATTATACATCGCTCCGGAGGTGCTTCTCCGCAAGGGTACACGAG	2364
O	y	2284	CTGTGTGACTGTGTGGAGTGTGGTGTATTCTTTGTGAAATGTTGGTGGGACAACTCCT	2343
D	b	2365	CTCTGTGACTGTGTGGAGCGTGGGTGTGATCTCTTTGAGATGCTGGTGGGACGCGCCT	2424
O	y	2344	TTCTTGGCACAAACCCATTAGAAACACACAAATGAAGGTTATCATCTGCGCAACTTCTTA	2403
D	b	2425	TTCTTGGCCCCCCCCACAGACGCGAGCTGGAAGGTGATCAACTGGGAGAGCACGCTG	2484
O	y	2404	CACATCCCTCCTCAAGCTAAGCTGAGTCTCTGAAGCCTCTGACCTCATTTATCAAACTGTGT	2463
D	b	2485	CATATCCTTACGAGGTGAGGCTCAGCGCTGAGGCCGAGACCTCATCCAGAGCTGTGC	2544
O	y	2464	CGAGGACCGAAGACCGCTTCGGCAAGAACGGTGTGTGAGATAAAGGCTCATCCATTT	2523
D	b	2545	TGCGCGCTGACTGCCCTGGGCGAGGATGGGCGAGATGACCTCAAGGCACACCCGCTC	2604
O	y	2524	TTTAAGACCATCATTTCTCTAGTGATCTGACACAGCAGTCTGCTTCATACATCCCTAAA	2583
D	b	2605	TTCAACACCATGACTTTTCCGCTGACATCCGAAAGAGGCTGCACCTAGTCCCCCACC	2664
O	y	2584	ATCAGCATCCAACAGATACATCCAATTTTCGACCCTGTTGATCCTGTGATAAATTTGTGGAGC	2643
D	b	2665	ATCAGCCACCCCATGGACACCTTCCAAATTTTGACCCGGTGGATGAAGAAAGCCCTTGGCAC	2724
O	y	2644	GATGGCAGCGGAGGAGAAATATCATGTGACACTCTGAGCGGATGATATAAAATTTGGGAAG	2703

Db 2725 GAGCCAGCGAGAG---AGCGCAAGGCTGGGACACGCTGGCCTCCCGCCAGCAGCAG 2781
 QY 2704 CACCCGAGCAGCCTTCTATGAGTTTCACTTTCGGAGGTTTTTTTGGATGACATGGCTAC 2763
 Db 2782 CTTCCAGAGCAGCCTTCTATGAGTTTCACTTTCGGAGGTTTTTTTGGATGACATGGCTAC 2841
 QY 2764 CCATATAATTATCCAAAGCCTATTGAGTATGAATACATT 2802.
 Db 2842 CCCTTCGGTCCCGAGGCTTCAGAGCCGCGAGAGT 2880

RESULT 11
 AAX87397
 ID AAX87397 standard; cdna; 5276 BP.
 AC AAX87397;
 DT 08-OCT-1999 (first entry)
 DE Human WART2 cdna.

WART2; hWART2; WART orthologue; human; signal transduction;
 protein kinase; cancer; tumour; diagnosis; therapy; ss.

Homo sapiens.
 Key Location/Qualifiers
 1..374
 5'UTR /*tag= a
 375..3641
 CDS /*tag= b
 3642..5276
 3'UTR /*tag= c

W09937787-A2.
 29-JUL-1999.
 20-JAN-1999; 99WO-US01145.
 21-JAN-1998; 98US-0072023.
 (SUGEN-) SUGEN INC.

Flanagan P, Plowman GD;
 WPI; 1999-458698/38.
 P-PSDB; AAY06527.

New nucleic acid encoding human orthologs of Drosophila WART
 proteins, used to identify specific modulators for treating cancer
 or for diagnosis

Claim 1; Page 122-123; 137pp; English.
 This is the nucleotide sequence of a cDNA clone coding for a human
 orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor
 serine/threonine kinase WART. hWART2 cDNA was isolated from a
 human bone marrow CDNA library using a PCR fragment of hWART1 as
 probe. hWART2 is consistently expressed in human tumour cells
 lines, except for most of the colon cancer lines examined. This
 overexpression in tumour cells versus normal tissues may provide a
 target for oncology drug development. Nucleic acids encoding
 full-length hWART2 and hWART2 polypeptides lacking one or more of
 amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518,
 and 974-1046, or lacking one or more of the N-terminal domain,
 catalytic domain, or C-terminal domain are claimed, as well as
 hWART1 sequences (AAX87396), hWART2 and hWART1 polypeptides,
 antibodies, a method for identifying modulators of hWART function,
 and use of such modulator compounds to treat an abnormal condition,
 involving hWART signal transduction, especially cancer. Probes for
 detection of hWART nucleic acids are also claimed.

XX

SQ Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;
 Query Match 24.5%; Score 787.2; DB 20; Length 5276;
 Best Local Similarity 68.7%; Pred. No. 9, 2e-186;
 Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;
 QY 1220 CAGTAAAGTAAAGCCCTGCAAGATGAACAGCTAGCTTTACCCAAAGAGATGATAGTGA 1279
 Db 1987 CAGGATGGAGCAGAGCTCCGTGGGGCCCAACGAGCGGAGGCGGACAGAGCC 2046
 QY 1280 AGAGTGGGACAGTGGTACTCTGGGATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1339
 Db 2047 GCAAAAGCGCCAAAGGGGACAAAGCGGAGGATTAAGAAAGAAAGAAAGAAAGAAAG 2106
 QY 1340 TCACGTCTTCGGAAGAAAGATGAAGACGAGAGAGTCTCGGATTCAGAGTTTACT 1399
 Db 2107 TTCCCGTCCGCAAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2166
 QY 1400 CCCCACAGCCCTTTAAGTTCTTCATGGAGCAGCAGTGTAGAGACGCTCTGAGTCTCATC 1459
 Db 2167 CGCATACGCTTTAAGTTCTTCATGGAGCAGCAGTGTAGAGAGAGAGAGAGAGAGAGAG 2226
 QY 1460 AGCAGCTGTGATCGGAAGAGAGAGCTAGAAATGAATGATGCGGGTTGGATTATCTC 1519
 Db 2227 AGCAGAGGTTTACCGGAGGCTGCAGCTGGCAAGAAATGSCCAAGAGCTGGAGCTCTGT 2286
 QY 1520 AAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTACTATATTCGTG 1579
 Db 2287 AGCTGAGCAGGAGCAGATGCGGAAGATCTCTACCAGAAAGAGTCTAATTTACAACAGT 2346
 QY 1580 TTAAGAGGCTTAAATGGACAGTCTATGTTTGAAGATTAAGACATTTAGGAATAGAG 1639
 Db 2347 TAAAGAGGCGCAAGATGGACAGTCTATGTTTGAAGATTAAGACATTTAGGAATAGAG 2406
 QY 1640 CGTTTGTGAAGTCTGTCTAGCAAGAAAGTCTGATCTACTACTACTACTACTACTACT 1699
 Db 2407 CCTTGGAGAGTGTGCTTGTGTAAGTGGACACTACGCGCTGTACGCCATGAAGA 2466
 QY 1700 CTCTCGAAAGAAAGAGTCTCTCGGAATCAGTGGCTCATGTGAAGCGGAGAGG 1759
 Db 2467 CCCTAAGGAAAGAGATGCTCTGAACCGGAATCAGTGGGCGCCACGTCAGGCGGAGAGG 2526
 QY 1760 ATATCTAGCAGAGGCGGACAAATGAGTGGTGGTCCGCCCTGTACTACTCTTTCCAGACA 1819
 Db 2527 ACATCTGGCGGAGGAGAGACAAATGAGTGGTGGTCAAACTCTACTCTCTCCAAAGACA 2586
 QY 1820 AGGACAACTTGTACTTGTGATGGACTACATTTCTGGGGGGATATGATGAGCTTATAA 1879
 Db 2587 AAGACAGCTGTACTTGTGATGGACTACATTTCTGGGGGGATATGATGAGCTTATAA 2646
 QY 1880 TTGAATGGGATCTTTCCTGAAATCTGGCAGATTTCTACATAGCAGAACTTACCTGTG 1939
 Db 2647 TCCGGATGGAGGTCTTCCCTGAGCAGCTGGCCCGGTTCTACATCGCAGAGCTGACTTGG 2706
 QY 1940 CAGTTGAAGTGTCTCAAAATGGGTTTTTATCATAGAGATTTAAACCTGTATACATTT 1999
 Db 2707 CCATTTGAGTGTCCCAAGATGGCTTTCATCCAGCAGAGATCAAGCTGTATACATTT 2766
 QY 2000 TGATTGACCGTGTGGCCATATTAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2059
 Db 2767 TGATAGATCTGGATGGTGCATTAATTAACCTACAGATTTCCGGCTCTGCACCTGGTTCAG 2826
 QY 2060 GGACACATGACTTCCAAAGTACTACAGAGTGGGATCACCCAGGCAAGATAGCATGATTT 2119
 Db 2827 GGATTCACAAATTCACAAATATTACCAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2886
 QY 2120 TCAGTAACGAATGGGGAGATCTCTCCATTTGCGTGGGAGCAGAGTGAAGCCACTGG 2179
 Db 2887 CCAGCGACCTCTGGGATGATGTGTCTAATCTGTGGTGGGAGCAGGCTGAAGACCCTAG 2946
 QY 2180 AGCGGAGAGCTGCTCGCCAGCAGCAGCAGGATGTCTAGCCCATTTCTGTTGGGACTCCA 2239
 Db 2947 AGCAGAGGCGCGGAGAGCAGCAGAGAGTGTCTGGCAGATTTCACTGCTGGGAGCTCAA 3006

QY 2240 ATTATATTGACCTGAAGTGTCTACTCGAAGGATATACACAGCTGTGTGCTGCTGGA 2299
 Db 3007 ACTACATCGCAGGCTGTCTCTCGAAGGATATACACAGCTGTGTGCTGCTGGA 3066
 QY 2300 GTGTTGGTGTATCTTGTGAATGTGTGGGACACCTCTTCTTGGCACAACCC 2359
 Db 3067 GTGTTGGTGTATCTTGTGAATGTGTGGGACACCTCTTCTTGGCACAACCC 3126
 QY 2360 CATTAGAAACAAATGAAGTGTATCATCTGCAACCTCTACACATCCCTCTCTCAAG 2419
 Db 3127 CCACAGAACCCAGCTGAAGTGTATCACTGGGAGAACACGCTCCACATTCAGCCCAAG 3186
 QY 2420 CTAAGCTGAGTCTGAAGCTCTGACCTCATATATCAAACTGTGTGAGGACACAGAGACC 2479
 Db 3187 TGAAGCTGAGCTGAGCCAGGACCTCATCAACAGCTGTGTGCTCTCCGACAGACACC 3246
 QY 2480 GCCTCGGCAAGACGGTGTATGAGATGAAGGCTATCACTTTTAAAGACCATCGATT 2539
 Db 3247 GCCTCGGCGGAATGGGCGGATGACCTGAAGGCCACCCCTTCTTCAAGCGCATGACT 3306
 QY 2540 TCTCTAGTGTCTGACAGAGCTGTCTCATATCCCTAAATCAGGATCCCAACAG 2599
 Db 3307 TCTCAGTGACATCCGAGGACGAGCCCTCTAGCTTCCACCATCAAGCCACCCATGG 3366
 QY 2600 ATACATCAATTTGACCTGTGTATCTGATATAAATTTGGAGCGATGGCAGCGAGGAGG 2659
 Db 3367 ACACCTCGAATTTGACCTGTGTATGAGAAAGCCCTTGGACATGCCAGC---GAAG 3423
 QY 2660 AAATATCAGTGACACTCTGAGCGGATGATATAAATTTGGAAGCACCACCGAGCGCTT 2719
 Db 3424 GTAGCACCAGGCTTGGACACACTCACCTCGCCCAATAACAAGCATCTGAGCAGCAT 3483
 QY 2720 TCTATGAGTTCACCTTTCCGAGGTTTGTATGACAAATGGCTACCCATATATATCAAA 2779
 Db 3484 TTTAGGAATTCACCTTCGGAAGGTTCTTTGATGACAAATGGCTACCCCTTTGATGCCCAA 3543
 QY 2780 AGCTATTGAGTATGATATCAATTTTACACAGGGCTCAGA 2819
 Db 3544 AGCTTCAGGACGACAGCTTCACAGGCTGAGAGCTCAGA 3583
 RESULT 12
 ID AAA59129 standard; DNA; 5486 BP.
 XX
 AC AAA59129;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS1.
 XX
 KW Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 387..3653
 FT /*tag= a
 FT /product= "tumour suppressor protein"
 XX
 PN EP1022333-A1.
 XX
 PD 26-JUL-2000.
 XX
 PF 07-OCT-1999; 99EP-0119199.
 XX
 PR 25-JAN-1999; 99JP-0016223.
 XX
 PA (JCRP-) JCR PHARM CO LTD.

XX Koga J, Kono K, Zolotaryov FN;
 XX WPI: 2000-516013/47.
 DR P-PSDB: AAB07633.
 XX
 PT New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 PT diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 PT nephropathy or cardiopathy
 XX
 PS Claim 1: Page 18-28; 59pp; English.
 XX
 CC The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnostic pharmaceutical preparations can be used for examining
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.
 XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;
 Query Match 24.5%; Score 787.2; DB 21; Length 5486;
 Best Local Similarity 68.7%; Pred. No. 9.4e-186;
 Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;
 QY 1220 CAGTAAGTAAAGCCCTGCAAGATGAACAGCTGTCTTACCAGGAGATGATAGTGA 1279
 Db 1999 CAGCATGAGCAGAGCCCTCGTGGGGCCCAACAGCCGCGGCGCACAGAGCC 2058
 QY 1280 AGAGTCGGACAGTGTGACTCTGGGATAAAGAAAGAACAGATTACACTTACACTA 1339
 Db 2059 GCAAAAGCGCCAAAGGGGACAAAGCGGAGGATAAAGAGCAGATTTCAGACCTCTCCG 2118
 QY 1340 TCACGTGTCGAAACAAAGAGATGAAGACGAGAGAGTCTCGGATTTCAGATTACT 1399
 Db 2119 TTCCCGTCCGAAACAAAGCAGAGAGAGAGAGAGTCTCGGATTTCAGATTACT 2178
 QY 1400 CCCACAGGCTTTAAGTTCCTTCATGGAGCAGCTAGAGAACCTCTCTGAAGTCTATC 1459
 Db 2179 CGCATACGCTTTAAGTTCCTTCATGGAGCAGCTAGAGAACCTCTCTGAAGTCTATC 2238
 QY 1460 AGCAGCGTCTGCATCGGAGAGAGAGCTAGAAATGAATGATGGGTTGATTTATCTC 1519
 Db 2239 AGCAGAGGTTAAGCGGAGGCTGCAGCTGGAGCAAGAAATGCGCAAGCTGGACTCTGTG 2298
 QY 1520 AAGATCCCGAGATCAATGAGAGATGCTTTCGCAAGAGCTCTAATCTATATTCGTC 1579
 Db 2299 AAGCTGACGAGGAGCAGATGCGAAGATCTCTTACGAAAGAGTCTAATTCACACAGT 2358
 QY 1580 TTAAGAGGCTTAAATGGACAGCTATGTTTGAAGATAAAGACATAGGAATAGGAG 1639
 Db 2359 TAAAGAGGCGCAAGATGACAGCTATGTTTGAAGATCAAAACCTGGGATCGGTG 2418
 QY 1640 CGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCTGATCTTAAAGCTTTGTATGCAACAA 1699
 Db 2419 CTTTGGAGAGTGTGCTTGTAAAGTGGACACTCAGCGCCCTCAGCCATGAAGA 2478
 QY 1700 CTCCTCGAAGAAAGACGCTTCTCGCAATAGTGGCTCATGTGAAAGCGAGAGGG 1759
 Db 2479 CCCTAAGGAAAGAGATGCTCTGAAACCGGATAGTGGCGCCAGCTCAAGCCGAGAGGG 2538
 QY 1760 ATATCTAGCAGAGCCGACAAATGAGTGGTGGTCCGCTCTACTACTCTTCTCCAGGAGA 1819
 Db 2539 ACATCTGCGCGAGCAGACAAATGAGTGGTGGTCCAAACTCTACTACTCTCTCCAGACA 2598
 QY 1820 AGGACAACTTGTACTTTGTGATGACTACATTCCTGGGGGATATGATGACCTATTAA 1879
 Db 2599 AAGACAGCCTGACTTTTGTGATGACTACATCCCTGGTGGGACATGATGACCTGCTGA 2658

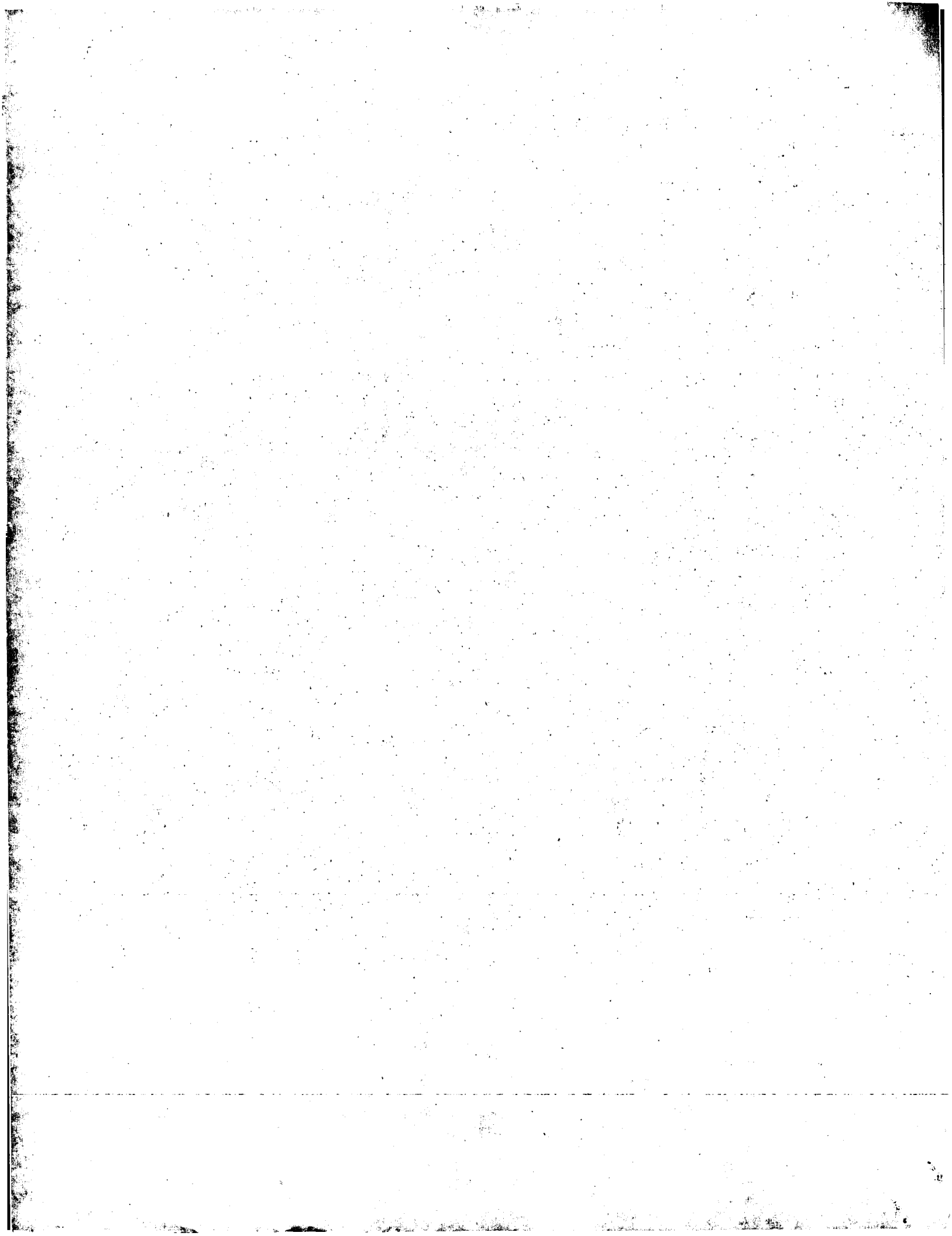
XX	DE	DNA encoding a tumour suppressor protein' hGHITS2.
XX	KW	Human; growth hormone inhibited tumour suppressor protein; hGHITS;
XX	KW	antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
XX	KW	diabetic nephropathy; cardiopathy; tumour; breast cancer;
XX	KW	renal adenocarcinoma; colorectal cancer; leukaemia; ss.
XX	OS	Homo sapiens.
XX	FH	Location/Qualifiers
XX	FT	CDS 387..3653
XX	ET	/**tag= a
XX	ET	/product= "tumour suppressor protein"
XX	PX	EPI022333-A1.
XX	PN	26-JUL-2000.
XX	PD	07-OCT-1999; 99EP-0119199.
XX	PF	25-JAN-1999; 99JP-0016223.
XX	PR	(JCRP-) JCR PHARM CO LTD.
XX	PA	Koga J, Kono K, Zolotaryov FN;
XX	PI	WPI: 2000-516013/47.
XX	DR	P-PSDB; AAB07664.
XX	DR	New human growth hormone inhibited tumour suppressor genes 1 and 2 for
XX	PT	diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
XX	PT	nephropathy or cardiopathy
XX	PT	Claim 1; Page 28-39; 59pp; English.
XX	PS	The present sequence encodes a human growth hormone inhibited tumour
XX	CC	suppressor (hGHITS) protein. The hGHITS proteins can be used in an
XX	CC	antineoplastic pharmaceutical preparation. Probes for the hGHITS
XX	CC	DNA sequences can be used in diagnostic pharmaceutical preparations.
XX	CC	The diagnostic pharmaceutical preparations can be used for examining
XX	CC	expression of hGHITS genes in dwarfism, gigantism, acromegaly,
XX	CC	angiopathy, diabetic nephropathy or cardiopathy, or in malignant
XX	CC	tumours including breast cancer, renal adenocarcinoma, colorectal
XX	CC	cancer, and leukaemia. Antibodies against the proteins can be used in
XX	CC	a diagnostic pharmaceutical preparation for examining expression of
XX	CC	a tumour suppressor gene.
XX	QY	Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;
XX	SQ	Query Match 24.5%; Score 787.2; DB 21; Length 5486;
XX		Best Local Similarity .687%; Pred. No. 9.4e-186;
XX		Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps
QY	1220	CAGTAAGTAGCCTCCAAGATGACACCGCTTACCACCANGAAGATCATAGTCAGA 1279
Db	1999	CAGCGATGAGCAGACGCCCTCCGTGGGGCCCCAACGAGCCGGGGCGGCACAAGAGCC 2058
QY	1280	AGAGTCGGCACAGTGGTGACTCTGGGGATAAGAAAACAAACAGATTACAACTTCACCTA 1339
Db	2059	GCAAAAGCCCAAGGGGGAACAAAGCGGAAGGATATAAAGCAGATTCAGACCTTCCCG 2118
QY	1340	TCACTCTTCGAAAAACAAGAAGATGAAGACGAGAGAGTCTCGGATTTCAGAGTTTACT 1399
Db	2119	TTCCCCGTCGCAAAACACGACGAGACGAAGAGAGAGATCAGCATCAAGAGCTTACT 2178
QY	1400	CCCACAGCGCCTTTAAGTTCTTCATGGAGCAGCAGCTAGAGAAAGCTCCTGAAGTCTCATC 1459
Db	2179	CGCCATACGCCCTTTAAGTTCTTCATGGAGCAGCAGCTGGAGAAATGTCATCAAACCTTACC 2238
QY	1460	ACGACGCTTCGATCGGAAGACGAGCTAGAAAATGAATATGCGGGTTGGATTATCTC 1519
Db	2239	ACGACGAAGTTAAACCGAGGCTGCAGCTGGAGCAAGAAATGCCAAAGCTTGACTCTGTG 2298

1

QY 1520 AAGATGCCAGGATCAATGAGAAAGATGCTTTGGCAGAAAGAGTCTAATATATTGSPC 1579
 DB 2299 AAGCTGACGACGAGCAGATGCGGAGATCCTCTACGAGAAAGAGTCTAATTAACACAGT 2358
 QY 1580 TTTAAAGGCTTAAATGGAAGTCTATGTTGTTAAAGATAAAGACATTTAGGAATAGGAG 1639
 DB 2359 TAAAGAGGCCAAGATGGCAAGTCTATGTTGTTCAAGATCAAAACCCCTGGGGATCGGTG 2418
 QY 1640 CTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATTAAGCTTTGTATGCAACAAGA 1699
 DB 2419 CTTTGGGAGAGTGTCTGTTTGAAGTGGACACCTACGCCCTGTACGCCATCAAGA 2478
 QY 1700 CTCCTCGAAAGAAAGAGCTTCTGCTCGGAAATCAGTGGCTCATGTGGAAGCGGAGAGG 1759
 DB 2479 CCTTAAGGAAAGAGATGTCCTGAACCGGAATCAGTGGCCCAAGTCAGGCGGAGAGG 2538
 QY 1760 ATATCTAGCAGAACCGCAATAGTGGTGGTCCGCTGTACTACTCTTTTCCAGGACA 1819
 DB 2539 ACATCTGCCCGAGGACAGACATGAGTGGTGGTCAAACTCTACTACTCTTCCAAAGACA 2598
 QY 1820 AGGACAATTTGACTTTGTGATGGACTACATTCCTGGGGGGATATGATGACCTATTAA 1879
 DB 2599 AAGACAGGCTGTACTTTGTGATGGACTACATTCCTGGTGGGACATGATGAGCCTGCTGA 2658
 QY 1880 TTAGAATGGGCATCTTCTGAAATCTGGCACGATTTCTACATAGCAGAACTTACTGTG 1939
 DB 2659 TCCGATGGAGTCTTCTGAGCAGCTGGCCGGTCTACATCGCAGAGCTGACTTTGG 2718
 QY 1940 CAGTTGAAAGTCTCATAAATGGGTTTATTATAGAGATATTAACCTGTATAACATTT 1999
 DB 2719 CCATTGAGAGTGTCCACAAGATGGGCTTCATCCCGGACATCAAGCCTGATACATTT 2778
 QY 2000 TGATTCACCTGATGCCATATTAATGACTGACTTTGGCTTGTGCACTGCTTCAGAT 2059
 DB 2779 TGATGATCTGGATGGTACATTAACCTACAGATTTGGGCTCTGCACTGGGTTCAGGT 2838
 QY 2060 GGACACATGACTTCCAAAGTACTACAGAGTGGGATCACCCAGCGAAGATAGCATGGAT 2119
 DB 2839 GGACTCACAAATTTCCAAATATTACAGAAAGGGAGCCATGTACAGACAGACACATGGAGC 2898
 QY 2120 TCAGTAACGAATGGGAGATCTTCCAAATGTCGGTGGGACAGACTGGAAGCCACTGG 2179
 DB 2899 CCAGCGAGCTCTGGGATGATGTCTTAAGTCTGCGTGGGACAGAGCTGGAAGACCTAG 2958
 QY 2180 AGCGAGAGCTCTCGCCAGCACACGAGTGTCTAGCCATTTCTGTTGGGACTCCCA 2239
 DB 2959 AGCAGAGGCGGAGGAGCAGCAGCAGAGGTGCTGCGACATTCATGTTGGGAGTCCAA 3018
 QY 2240 ATTATATTGCACTGAAGTGTCTTCCGAAACAGGATATACAGAGTGTGTGACTGTGGA 2299
 DB 3019 ACTACATGCAACCCAGGCTCTCTCCGCAAGGATACACTCACTGTGACTGGTGA 3078
 QY 2300 GTGTTGGTGTATCTTTGTGAAATGTTGGTGGGACAACTCTTCTTTGGCACAACCC 2359
 DB 3079 GTGTTGGAGTGTATCTCTCGAGATGCTGTGGGGCAGCGCCCTTTTGGCAGCTACTC 3138
 QY 2360 CATTAGAAACAAATGAAGGTTATCTCTGCAAACTTCTACACATCCCTCTCAAG 2419
 DB 3139 CCACAGAACCCAGCTGAAGGTGATCACTGGGAGACACGCTCCACATTCAGGCCAGG 3198
 QY 2420 CTAAGCTGAGTCTGAAGCCTTGACCTCATTTATCAAACTGTGTGGAGCAGCAAGACC 2479
 DB 3199 TGAAGCTGAGCCTGAGGCCAGGACCTCATCAACCAAGTGTGCTCCGAGCAGCACC 3258
 QY 2480 GCCTCGGAAGAACGGTGTGTATGAGATAAAGGCTCATCCATTTTAAAGACCATCGATT 2539
 DB 3259 GCCTGGGGCGGAATGGGCGGATGATACCTGAAGGCCACCCCTTCTTCAGCGCATGACT 3318
 QY 2540 TCTCTAGTGTATGACAGCAGAGTGTCTTATACATCCCTAAATCAGGCATCAACAG 2599
 DB 3319 TCTCAGTGTACATCCGGAAGCAGCAGCCGCCCTTACGTCTCCCACTACGACCCACCTG 3378

QY 2600 ATACATCCAAATTCGACCCCTGTGTGATCTCTGATATAAAATTTGTGAGCGATGCGACGAGG 2659
 DB 3379 ACACCTCGAATTCGACCCCTGTGTGATGAGAAGCCCTTGGACCATGCCAGC---GAG 3435
 QY 2660 AAAATATAGTACACTCTGACCGGATGGTATAAAATGGGAAGCACCCTGAGCAGCTTT 2719
 DB 3436 GTAGCACCAGCCCTGGGACACACTCACCTGCCCAATAACAAGCATCTGAGCAGCAT 3495
 QY 2720 TCTATGATTCACCTTCGGAGGTTTGTGATGACAATGGCTACCATATATATCCAA 2779
 DB 3496 TTTAGCAATTCACCTTCGGAAGGTTCTTTGATGACAATGGCTACCCCTTTCGATGCCCAA 3555
 QY 2780 AGCCTATTGAGTATGAATACATTCATTCACAGGGCTCAGA 2819
 DB 3556 AGCCTTCAGGACAGAGCTTCACAGGCTGAGAGCTCAGA 3595

RESULT 14
 AAZ61160
 ID AAZ61160 standard; cDNA; 1961 BP.
 XX
 AC AAZ61160;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE cDNA SS1771A encoding a partial protein kinase.
 XX
 KW Kinase activity; molecular weight marker; Isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..1769
 FT /*tag= a
 FT /note= "partial sequence"
 XX
 PN WO200008180-A2.
 XX
 PD 17-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17630.
 XX
 PR 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0095972.
 XX
 PA (IMMV) IMMUNEX CORP.
 XX
 PI Virca GD, Bird TA, Anderson DM, Marken JS;
 XX
 DR WPI: 2000-195584/17.
 DR P-PSDB; AAY69162.
 XX
 PT New human kinase polypeptides and polynucleotides used as molecular.
 PT weight markers and as controls for peptide fragmentation
 XX
 PS Claim 1; Page 9; 60pp; English.
 XX
 CC The present sequence encodes a partial polypeptide which has kinase
 CC activity. The kinase polynucleotides can be used to express the
 CC polypeptides, and as probes to identify nucleic acids encoding
 CC proteins having kinase activity. The kinase polypeptides and
 CC fragmented polypeptides are used as molecular weight and isoelectric
 CC focusing markers, and as controls for peptide fragmentation. They also
 CC have a number of therapeutic uses as kinases play a central role in
 CC cellular signal transduction. The polypeptides could also be used as
 CC identify binding partner proteins. The polypeptides can also be used as
 CC a reagent to identify any proteins that the polypeptide regulates, and
 CC proteins with which it might interact. The polypeptides may also be
 CC used for preparation of antibodies. The antibodies can be used in
 CC assays to detect the presence of the protein, and to purify the protein
 CC by immunoaffinity chromatography.
 XX



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:02 ; Search time 65.3715 Seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtcgaacattcaattacgcg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2.6/ptodata/1/ina/PTC05_COMB.seq:*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3213	100.0	3213	4	US-09-442-100-5
2	2483	77.3	3984	4	US-09-442-100-3
3	812.6	25.3	3155	4	US-09-442-100-7
4	785.6	24.5	1961	4	US-09-509-902A-15
5	713.8	22.2	1498	4	US-09-509-902A-6
6	521.2	16.2	5720	4	US-09-442-100-1
7	518.6	16.1	638	4	US-09-328-111-26
8	376.6	11.7	678	4	US-09-328-111-66
9	248.4	7.7	1935	2	US-08-878-989-11
10	248.4	7.7	1935	4	US-09-272-796-11
11	244.4	7.6	3018	2	US-08-860-150-6
12	244.4	7.6	3018	3	US-09-338-132-6
13	198.6	6.2	2101	2	US-08-860-150-1
14	198.6	6.2	2101	3	US-09-338-132-1
15	179.8	5.6	2160	4	US-09-588-256-1
16	125.4	3.9	4739	3	US-08-685-871-1
17	115.8	3.6	2311	2	US-08-712-709-6
18	115.8	3.6	2311	3	US-09-111-444-6
19	115.8	3.6	2311	4	US-09-541-228-6
20	115.8	3.6	2370	4	US-09-031-295-1
21	115.4	3.6	5053	3	US-08-685-576-2
22	112.8	3.5	1599	3	US-09-256-465-1
23	112.8	3.5	1599	4	US-09-167-322-3
24	109	3.4	4363	2	US-08-685-576-5
25	107	3.3	2549	4	US-09-467-082-3
26	105	3.3	2726	2	US-08-422-699A-12
27	105	3.3	2726	2	US-08-422-706B-12

28	105	3.3	3182	1	US-08-484-044-11
29	103.4	3.2	2511	2	US-08-422-699A-8
30	103.4	3.2	2511	2	US-08-422-706B-8
31	97.2	3.0	2706	2	US-08-630-822A-61
32	97.2	3.0	2706	2	US-09-005-069-61
33	97.2	3.0	2706	4	US-09-171-156A-20
34	96.4	3.0	2104	4	US-09-313-930-1
35	96.4	3.0	2196	1	US-08-313-274-1
36	96	3.0	2244	3	US-09-094-714A-48
37	96	3.0	2245	4	US-09-225-749-24
38	88.2	2.7	2754	4	US-09-429-322-3
39	78.8	2.5	1732	4	US-09-430-564-1
40	78.2	2.4	1890	3	US-09-289-466-2
41	78.2	2.4	1929	2	US-09-016-000-10
42	77.6	2.4	2599	6	5266464-1
43	76.8	2.4	2557	4	US-08-464-954A-1
44	76.6	2.4	1891	3	US-09-289-466-1
45	73.6	2.3	2610	2	US-09-212-771-1

ALIGNMENTS

RESULT 1
US-09-442-100-5
; Sequence 5, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3213 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2889

US-09-442-100-5

Query Match:		100.0%;	Score 3213;	DB 4;	Length 3213;			
Best Local Similarity		100.0%;	Pred. No. 0;					
Matches 3213;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
Qy	1	GTGCAACATTC	AAATTAACCG	AAACAAAGCT	TGGAAGTTC	TCTAAGAGTCT	CTAGTTCCT	60
Db	1	GTGCAACATTC	AAATTAACCG	AAACAAAGCT	TGGAAGTTC	TCTAAGAGTCT	CTAGTTCCT	60
Qy	61	CAGAGACACG	CCCATCTCT	AGGAGAAAT	TGCTTTAT	CGTTTATG	CTTGAAGCCCACTCA	120
Db	61	CAGAGACACG	CCCATCTCT	AGGAGAAAT	TGCTTTAT	CGTTTATG	CTTGAAGCCCACTCA	120
Qy	121	CAGCGGATG	TAGGAAGAC	CTCTGT	CTGGATCG	GGCATTTG	CTCAGCATTTG	180
Db	121	CAGCGGATG	TAGGAAGAC	CTCTGT	CTGGATCG	GGCATTTG	CTCAGCATTTG	180
Qy	181	CCAGCAATG	AGAGAGTGA	ACCCCTC	CCACACAC	CTCAAGT	TAGGAGTCT	240
Db	181	CCAGCAATG	AGAGAGTGA	ACCCCTC	CCACACAC	CTCAAGT	TAGGAGTCT	240
Qy	241	CCACCACTC	CGAGAGCC	CAGACCC	ACCTTCC	CCGAGG	CACCACTCC	300
Db	241	CCACCACTC	CGAGAGCC	CAGACCC	ACCTTCC	CCGAGG	CACCACTCC	300
Qy	301	TGGGAACCA	AGCTCTC	CAGACA	AGCGCTA	CTCTGG	GAACATGG	360
Db	301	TGGGAACCA	AGCTCTC	CAGACA	AGCGCTA	CTCTGG	GAACATGG	360
Qy	361	ATCTCCCTG	TTCACCTC	GGGCTGG	CAGAGGG	GTACCTT	CCACACCTT	420
Db	361	ATCTCCCTG	TTCACCTC	GGGCTGG	CAGAGGG	GTACCTT	CCACACCTT	420
Qy	421	TCTCCCATG	ATCCCTT	AGCAGG	CTCAG	AGGGCC	ATTAGTTC	480
Db	421	TCTCCCATG	ATCCCTT	AGCAGG	CTCAG	AGGGCC	ATTAGTTC	480
Qy	481	CAACCATAT	CATCAGAG	TACTAG	CAAA	TCTTAC	ACAGGCG	540
Db	481	CAACCATAT	CATCAGAG	TACTAG	CAAA	TCTTAC	ACAGGCG	540
Qy	541	CAGAATGGT	GTGGTCT	AGTCTG	ATTTAT	CGTGCA	CAAAATG	600
Db	541	CAGAATGGT	GTGGTCT	AGTCTG	ATTTAT	CGTGCA	CAAAATG	600
Qy	601	ACTCGGACG	CCACCTC	CTATCT	CTGAC	CCCAAA	TGTCAC	660
Db	601	ACTCGGACG	CCACCTC	CTATCT	CTGAC	CCCAAA	TGTCAC	660
Qy	661	TTACAAACG	GGGCTTCT	GTCTG	CTCCAC	CAATAT	TGCGCA	720
Db	661	TTACAAACG	GGGCTTCT	GTCTG	CTCCAC	CAATAT	TGCGCA	720
Qy	721	ATGATGGTG	CCCAACAG	CAAGTCA	TAAATG	AGGCTTT	ATAATG	780
Db	721	ATGATGGTG	CCCAACAG	CAAGTCA	TAAATG	AGGCTTT	ATAATG	780
Qy	781	CTGCAACAG	CGCTGG	CCCGCTG	CTTCTG	CTCGC	AGTCA	840
Db	781	CTGCAACAG	CGCTGG	CCCGCTG	CTTCTG	CTCGC	AGTCA	840
Qy	841	CATGAATTC	CTACATG	GCACCTT	ACCATG	AGGTC	CAAAATG	900
Db	841	CATGAATTC	CTACATG	GCACCTT	ACCATG	AGGTC	CAAAATG	900
Qy	901	TTAGGAAGT	AGAGCA	AGTCTG	CTAATC	TCTCAG	CGCTT	960
Db	901	TTAGGAAGT	AGAGCA	AGTCTG	CTAATC	TCTCAG	CGCTT	960
Qy	961	ATCACACCC	CGCTCCT	ATTTCA	ACAGCCG	TGAAAG	CATGCG	1020
Db	961	ATCACACCC	CGCTCCT	ATTTCA	ACAGCCG	TGAAAG	CATGCG	1020

Qy	1021	CAGACTGCT	TTAGCCCC	CAACCATC	CTTCTT	GGATGCC	ACAGCC	AGTTCA	GACTGTT	CAG	1080
Db	1021	CAGACTGCT	TTAGCCCC	CAACCATC	CTTCTT	GGATGCC	ACAGCC	AGTTCA	GACTGTT	CAG	1080
Qy	1081	CCTACCCCT	TTTTCTT	CTGAGG	TACAGT	TCAAGT	GTGCT	GTATCCC	ACCTGTT	GCTGAA	1140
Db	1081	CCTACCCCT	TTTTCTT	CTGAGG	TACAGT	TCAAGT	GTGCT	GTATCCC	ACCTGTT	GCTGAA	1140
Qy	1141	GCTCCAGCT	ATCAAGT	TCCAC	CGCTTAT	CCAAAC	ATCTG	CTACAC	CAAAAC	CCCA	1200
Db	1141	GCTCCAGCT	ATCAAGT	TCCAC	CGCTTAT	CCAAAC	ATCTG	CTACAC	CAAAAC	CCCA	1200
Qy	1201	TCTGTCCCT	CCATATG	ATGATC	AGTAA	AGCCCT	TGCAAG	ATGAAC	AGCCCT	TAGCTT	1260
Db	1201	TCTGTCCCT	CCATATG	ATGATC	AGTAA	AGCCCT	TGCAAG	ATGAAC	AGCCCT	TAGCTT	1260
Qy	1261	AAGAAGAT	GATAGT	GAGAAG	AGTGG	CACAGT	GTGCT	GACTCT	GGGGAT	AAAGAA	1320
Db	1261	AAGAAGAT	GATAGT	GAGAAG	AGTGG	CACAGT	GTGCT	GACTCT	GGGGAT	AAAGAA	1320
Qy	1321	CAGATTACA	CTTCA	CTTATC	CTTCT	CGGAAA	CAAG	AAAGAT	GAAG	CAAGAG	1380
Db	1321	CAGATTACA	CTTCA	CTTATC	CTTCT	CGGAAA	CAAG	AAAGAT	GAAG	CAAGAG	1380
Qy	1381	TCTCGATTC	CAGATT	TACTC	CCAC	AGCCCT	TAAAGT	TTCTAT	GGAG	CAGC	1440
Db	1381	TCTCGATTC	CAGATT	TACTC	CCAC	AGCCCT	TAAAGT	TTCTAT	GGAG	CAGC	1440
Qy	1441	AACCTCTG	AAGTCT	CATC	AGCAG	CGCTC	ATC	CGGAA	AGCAG	CTAG	1500
Db	1441	AACCTCTG	AAGTCT	CATC	AGCAG	CGCTC	ATC	CGGAA	AGCAG	CTAG	1500
Qy	1501	ATCGGGTTC	GGATAT	CTCA	AGATG	CCCA	AGAT	CAAT	GAGA	AGAT	1560
Db	1501	ATCGGGTTC	GGATAT	CTCA	AGATG	CCCA	AGAT	CAAT	GAGA	AGAT	1560
Qy	1561	GAGTCTAAC	TATAT	TCGTCT	TAAAG	GGCTT	AAATG	CGCA	AGTCT	ATGTT	1620
Db	1561	GAGTCTAAC	TATAT	TCGTCT	TAAAG	GGCTT	AAATG	CGCA	AGTCT	ATGTT	1620
Qy	1621	AAGACAT	TAGGA	TAGAG	GGTGG	TGTAAG	TCTGT	CTAG	CAAG	AAAG	1680
Db	1621	AAGACAT	TAGGA	TAGAG	GGTGG	TGTAAG	TCTGT	CTAG	CAAG	AAAG	1680
Qy	1681	GCCTTCT	TATG	CAAC	AAAG	ACTCT	TCG	AAAG	AAAG	AGCT	1740
Db	1681	GCCTTCT	TATG	CAAC	AAAG	ACTCT	TCG	AAAG	AAAG	AGCT	1740
Qy	1741	CATGTGA	AGCGG	AGAGG	GATAT	CTCT	AG	CAAG	CGG	CAAT	1800
Db	1741	CATGTGA	AGCGG	AGAGG	GATAT	CTCT	AG	CAAG	CGG	CAAT	1800
Qy	1801	TACTACTCT	TTCC	AGGAC	AGCA	ACTTGT	TACT	TTGT	GATG	AGTAC	1860
Db	1801	TACTACTCT	TTCC	AGGAC	AGCA	ACTTGT	TACT	TTGT	GATG	AGTAC	1860
Qy	1861	GATATGAT	GAGCCT	TATTA	TAG	ATG	GGCAT	CTTCT	CTG	AAAT	1920
Db	1861	GATATGAT	GAGCCT	TATTA	TAG	ATG	GGCAT	CTTCT	CTG	AAAT	1920
Qy	1921	ATAGCAG	AACTT	ACCTG	TCG	AGT	TGA	AGT	TTCA	AAAT	1980
Db	1921	ATAGCAG	AACTT	ACCTG	TCG	AGT	TGA	AGT	TTCA	AAAT	1980
Qy	1981	ATTAACCT	GTAT	AACAT	TTTTC	AT	TG	CACT	GTAT	TAAT	2040
Db	1981	ATTAACCT	GTAT	AACAT	TTTTC	AT	TG	CACT	GTAT	TAAT	2040
Qy	2041	TTGTGCA	CTG	CTT	CAG	ATG	GAC	ACAT	GACT	TCC	2100
Db	2041	TTGTGCA	CTG	CTT	CAG	ATG	GAC	ACAT	GACT	TCC	2100

QY 2101 CGCAGATAGCATGATTTTCAGTACGAATGGGAGATCCTTCCAAATGTGCGGTGGG 2160
DB 2101 CGCAGATAGCATGATTTTCAGTACGAATGGGAGATCCTTCCAAATGTGCGGTGGG 2160
QY 2161 GACAGACTGAAGCCACTGAGCGGAGAGTGTCTGCCACACAGCGATGTCTAGCCAT 2220
DB 2161 GACAGACTGAAGCCACTGAGCGGAGAGTGTCTGCCACACAGCGATGTCTAGCCAT 2220
QY 2221 TCTCTGTTGGGACTCCCAATATATTGACCTGAAGTCTACTGCGAAGAGATATACA 2280
DB 2221 TCTCTGTTGGGACTCCCAATATATTGACCTGAAGTCTACTGCGAAGAGATATACA 2280
QY 2281 CAGCTGTGTGACTGGTGGAGTGTGGTGTATTTTGTGAATGTGTGGGACAACCT 2340
DB 2281 CAGCTGTGTGACTGGTGGAGTGTGGTGTATTTTGTGAATGTGTGGGACAACCT 2340
QY 2341 CTTTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAACTTCT 2400
DB 2341 CTTTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAACTTCT 2400
QY 2401 CTACACATCCCTCCTCAAGCTAAGTGTAGTCTGAAGCTCTGACCTCATTTCAACATG 2460
DB 2401 CTACACATCCCTCCTCAAGCTAAGTGTAGTCTGAAGCTCTGACCTCATTTCAACATG 2460
QY 2461 TGTGAGGACCAAGAGACCGCTCGSCAAGACGGTGTGTGATGAGATAAGCTCATCCA 2520
DB 2461 TGTGAGGACCAAGAGACCGCTCGSCAAGACGGTGTGTGATGAGATAAGCTCATCCA 2520
QY 2521 TTTTAAAGACCATGATTTCTTAGTGTATGTGACAGCAGTCTGCTTCATACATCCCT 2580
DB 2521 TTTTAAAGACCATGATTTCTTAGTGTATGTGACAGCAGTCTGCTTCATACATCCCT 2580
QY 2581 AAAATCACCATCCCAACAGATACATCCAAATTTGACCTGTGTGATCCTGATAAATGTGG 2640
DB 2581 AAAATCACCATCCCAACAGATACATCCAAATTTGACCTGTGTGATCCTGATAAATGTGG 2640
QY 2641 AGCGATGGCAGCGAGGAGGAAAATATCAGTGACACTCTGAGCGGATGGTATAAAATGGG 2700
DB 2641 AGCGATGGCAGCGAGGAGGAAAATATCAGTGACACTCTGAGCGGATGGTATAAAATGGG 2700
QY 2701 AAGCAGCCGAGCAGCGCTTCTATGATGTTCACTTTCGAGGTTTTTGTGATGACAAATGGC 2760
DB 2701 AAGCAGCCGAGCAGCGCTTCTATGATGTTCACTTTCGAGGTTTTTGTGATGACAAATGGC 2760
QY 2761 TACCATATAATTTATCCAAAGCTTATGATGATGAATACATTCATTTCAGGCGCTCAGAA 2820
DB 2761 TACCATATAATTTATCCAAAGCTTATGATGATGAATACATTCATTTCAGGCGCTCAGAA 2820
QY 2821 CAACAGTCTGATGAAGATGATCAACACACAAAGCTCCGATGGAACACCGAGATCTAGTG 2880
DB 2821 CAACAGTCTGATGAAGATGATCAACACACAAAGCTCCGATGGAACACCGAGATCTAGTG 2880
QY 2881 TATGTTTAAATTAACCTAGGAGATCATTTGAAGATTTGCAAGAGGCTGAAGTGCAGGGGT 2940
DB 2881 TATGTTTAAATTAACCTAGGAGATCATTTGAAGATTTGCAAGAGGCTGAAGTGCAGGGGT 2940
QY 2941 TTTTGAAGTTTGAAGAAATATGCAAAATGTCACAGAGTTTCTGTCTGTGTGTACAATA 3000
DB 2941 TTTTGAAGTTTGAAGAAATATGCAAAATGTCACAGAGTTTCTGTCTGTGTGTACAATA 3000
QY 3001 TTTTATTTTCCCTAAGTATGGAATTTGTTTAAATGTTAATTTATCCACCCCTTTTAA 3060
DB 3001 TTTTATTTTCCCTAAGTATGGAATTTGTTTAAATGTTAATTTATCCACCCCTTTTAA 3060
QY 3061 TTCAGTAAATTTAGAAAAATTTGTTATAGGAAGTAAATTTATGACTGAGTATTATAGTC 3120
DB 3061 TTCAGTAAATTTAGAAAAATTTGTTATAGGAAGTAAATTTATGACTGAGTATTATAGTC 3120
QY 3121 AATTTCTGTACTTAAAGTACTTAAAGAGAGAGCTGGTATCTTTTGTATATATAATAA 3180
DB 3121 AATTTCTGTACTTAAAGTACTTAAAGAGAGAGCTGGTATCTTTTGTATATATAATAA 3180
QY 3181 ATAATTTTAAATCCCAAAAAA 3213

DB 3181 ATAATTTTAAATCCCAAAAAA 3213
RESULT 2
US-09-442-100-3
; Sequence 3, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Welyi
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATIS
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 231..3623
; US-09-442-100-3
Query Match 77.3%; Score 2483; DB 4; Length 3984;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
QY 1 GTGCAACATTCATTAACCGAAACAAAGAGTCTGAAAGAGTCTTAAAGAGTCTCTAGTCTCT 60
DB 732 GTGCAACATTCAGTTAAACCGAAACAAAGAGTCTGAAAGAGTCTTAAAGAGTCTCTAGTCTCT 791
QY 61 CAGACAGCGGCGCCATCTCTAGGAGAAAATGTGGTGTATCTGCTGTAAGAGCCCAACTCA 120
DB 792 CAGAGGCATGGCCCGCCACTAGGAGAAAATGTGGGCTATCATCTCAGAGTCCCAACTCA 851
QY 121 CAGCGCGATGTAGGAGACCTCTGCTGATCCCGCATTCGAGCATTTCTCAAGCTCAC 180
DB 852 CAGACAGATGTAGGAGACCTTTGTGCTGATCTGTGATATATCAGCATTTGTTCAGAGCTCAC 911
QY 181 CCAAGCAATGGACAGAGAGTGAACCCCGCCCAAGCTTCAAGTTAGGAGTGTCTACTCT 240

Db 912 CCTAGCAACGACAGAGAGTGAACCCGCCACCACCACTCAAGTAAGGAGTGTTACTCCT 971
QY 241 CCACCACCTCCGAGAGGCGACGCCACCTCCCGGAGGACACACTCCCTCCGCCCTCA 300
Db 972 CCACCACCTCCGAGAGGCGAGACTCCCTCCGAGGAGTACAACTCCACCTCCCTCC 1031
QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCGA 360
Db 1032 TGGGAACCAACACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAATCTCCGA 1091
QY 361 ATCTCCCTGTTCCACTCGGGGTGGGAGGAGGAGTACCTCCACCACTCTTACCAC 420
Db 1092 ATCTCTCTGTCCTCCACTCGGGCATGGCAAGAGGCTATCTCTCCACCACTCTCAAC 1151
QY 421 TCTCCCATGAATCCCTTAGCAGGCTCAGAGGCGCTAGTCTGTTCCAGTGGTAGA 480
Db 1152 TCCCCCATGATCTCTCAACAGAGGAGGAGGAGTCTGTTCTCTGTTGGCAGA 1211
QY 481 CAACCCATCATGAGAGTACTAGCAAAATTTAACTTTTACACAGGCGGACCTGGAGTT 540
Db 1212 CAACCAATCATGAGAGTCTAGCAAAATTTAACTTTTCCATCAGGAGGAGCTGGAATG 1271
QY 541 CAGAAATGGTGGTCTGATTTATCTGTCGACCAAAA---TGTCCCACTGGTCT 597
Db 1272 CAGAAATGGTGGGCAAACTGATTTATGATACACCAAAATGTTGTCCTGCTGGCACT 1331
QY 598 GTGACTCGGACGCCACCTCCATATCTCTGACCCAGCTTAATGGCAAGGCCCTCT 657
Db 1332 GTGAATCGGACGCCACCTCCATATCTCTGACAGCAGCTAATGGCAAGGCCCTCT 1391
QY 658 GTTTTAAACAGGGGCTTCTGCTCTCCACATCATTCGCAATGGAAGCTTCTCTCAG 717
Db 1392 GTTTTAAACAGGGGAGTCTGCTCTCTGTCATATACAAATGGAAGTATCTCTCAG 1451
QY 718 TCGATGATGGTCCCAACAGGAACAGTCTATACATGAGGAGCTTAAATTAATGTCCT 777
Db 1452 TCTATGATGGTCCCAACAGGAATAGTCTATACATGGAAGTATTAATTAATGTCCT 1511
QY 778 GAGTGCACAGGCTGCGCCAGTCTGCTCTGCTGCGAGTCAATCCCAAGCGGT 837
Db 1512 GAGTGCACAAACAAATGCGCTCAGTCTCTGCTGCGAGCCAGTCAATCCCAAGCGGT 1571
QY 838 GGGCATGAAATTCCTACATGCACTCAATACCACTGAGTCAATCTCTTTAAATAC 897
Db 1572 GGGCATGAAATTCCTACATGCACTCAATACCACTGAGTCAATCTCTTTAAATAC 1631
QY 898 CCATTAGGAAGTAGAGAGTCACTCTGCTAATCTCTGAGCTTCTGCACTACAGTCACT 957
Db 1632 CCATTAGGAATAGAGAGTCACTCTGCTAATCTCTGAGCTTCTGCACTACAGTCACT 1691
QY 958 GCCATCACCGCTCTTATTCACAGCCCTGAAAGCAATGCGGCTGCTGAAACCAAGAG 1017
Db 1692 GCAATACACAGCTCTTATTCACAGCCCTGAAAGTATGCGTGTATTAACCAAGAG 1751
QY 1018 CTGACAGTCTGTTAGGCCCAACCACTCTCTTGGATGCAAGCCAGTCACTGTT 1077
Db 1752 CTACAGTCTGTTAGCACCTACACCCCTCTTGGATACCAAGCCAACTCAAACTGTT 1811
QY 1078 CAGCCTACCCCTTTCTGAGGGTACAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCT 1137
Db 1812 CAACCCAGTCTCTTCTGAGGAAACCGCTTCAATGTGATGCTGCTGCTGCTGCTGCT 1871
QY 1138 GAAGCTCCAGCTATCAAGTTCACAGCCGCTTATCCAAACATCTGCTACACCAAAAC 1197
Db 1872 GAAGCTCCAAACTATCAAGGACCAACCACTTACCCCAAACTGCTGCTGCTGCTGCT 1931
QY 1198 CCATCTGCTCCCTCATGATGAGTCAAGTAAAGCCCTGCAAGATGAACAGCTAGCTTA 1257
Db 1932 CCATCTGCTCTCATGAGTCAATCAGTAAAGCTGCAAGAGGATCAAGCAAGCTTG 1991
QY 1258 CCNAGGAAGATGATGAGAGAGTGGGAGAGTGGTGAAGTCTGGGGAATAAGAAAG 1317

Db 1992 CCCAAGGAGATGACAGTGAAGAGAGTATGAANAATGTTAGTGGGATAAAGAAAG 2051
QY 1318 AACAGATTTACAACCTTCACTCTTACCTTTCGGAAAAACAAGAGATGAAGAACGAAGA 1377
Db 2052 AACAGATTTACAACCTTCACTCTTACCTTTCGGAAAAACAAGAGATGAAGAACGAAGA 2111
QY 1378 GAGTCTCGATTCAGAGTTACTCCCAAGAGGCTTTAAAGTCTTCTATGAGAGCAGCAGTA 1437
Db 2112 GAATCTCGATTTCAAAAGTTATCTCTCAAGCAATTTAAATTTCTTTATGAGCAACATGA 2171
QY 1438 GAGAACGCTCTGAAGTCTCATCAGCAGCTCTGATCGGAAGAGAGCTGAGAAATGAA 1497
-Db 2172 GAAATGCTACTCAATCTCATCAGCAGCTCTACATCTGTAANAACAAATAGAGAAATGAA 2231
QY 1498 ATGATGCGGGTGGATATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAG 1557
Db 2232 ATGATGCGGGTGGATATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAA 2291
QY 1558 AAGAGTCTAACTATATTTCGCTTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 1617
Db 2292 AAGAACTTAATTAATCTCCGCTTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 2351
QY 1618 ATAAAGACATTTAGAAATAGGAGGCTTTGGTGAAGTCTCTAGCAAGAAAGTGCATACT 1677
Db 2352 ATAAAGACATTTAGAAATAGGAGCTTTGGTGAAGTCTCTAGCAAGAAAGTGCATACT 2411
QY 1678 AAGCTTTGATGCAACAAAGACTTTCGAAAGAAAGAGCTTCTGCTCCGAAATFAGGTTG 1737
Db 2412 AAGCTTTGATGCAACAAAGACTTTCGAAAGAAAGAGCTTCTGCTCCGAAATFAGGTTG 1797
QY 1738 GCTCATGTGAAGGCGAGAGGATATCTTAGCAAGACCCGACAAATGAGTGGTGGTCCGC 1797
Db 2472 GCTCATGTGAAGGCGAGAGGATATCTTAGCAAGACCCGACAAATGAGTGGTGGTCCGC 2531
QY 1798 CTGTACTACTCTTCCAGGACAGGACAACTTGTACTTTGTGATGAGTCACTTCTCTGGG 1857
Db 2532 CTATATTATTCATTCACAAAGATAAGGACAAATTTACTTTGTAATGAGTCACTTCTCTGGG 2591
QY 1858 GGGGATATGATGAGGCTTAAATTTAGAAATGAGTCTTCTGCAAAATCTGCGACCAATTC 1917
Db 2592 GGTGATATGATGAGGCTTAAATTTAGAAATGAGTCTTCTGCAAAATCTGCGACCAATTC 2651
QY 1918 TACATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGGTTTTTATTCATAGA 1977
Db 2652 TACATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGGTTTTTATTCATAGA 2711
QY 1978 GATATTAAACCTGATACATTTTGTGAGCTGATGGCCATATTAATGAGTCACTT 2037
Db 2712 GATATTAAACCTGATACATTTTGTGAGTGAAGTGTTCATAAAATGGGTTTTTATTCATAGA 2771
QY 2038 GGCTTGTGACTGGCTTCAGATGACACATGACTCCCAAGTACTACAGAGTGGGATCAC 2097
Db 2772 GGCTTGTGACTGGCTTCAGATGACACATGACTCCCAAGTACTACAGAGTGGGATCAC 2831
QY 2098 CCAGGCAAGATGAGTGGATTCAGTAACGAATGGGAGATCTTCCAAATTTCCGGTGT 2157
Db 2832 CCAGGCAAGATGAGTGGATTCAGTAACGAATGGGAGATCTTCCAAATTTCCGGTGT 2891
QY 2158 GGGGACAGACTGAAGCCACTTGGAGCGGAGCTGCTGCGCCAGCAGCAGGATGCTTAGCC 2217
Db 2892 GGNACAGACTGAAGCCACTTGGAGCGGAGCTGCTGCGCCAGCAGCAGGATGCTTAGCA 2951
QY 2218 CATCTCTGTTGGGACTCCCAATTTATTTGACCTGAAAGTGTCTACTGGAACAGGATAT 2277
Db 2952 CATCTCTGTTGGGACTCCCAATTTATTTGACCTGAAAGTGTCTACTGGAACAGGATAT 3011
QY 2278 ACACAGCTGTGACTGGTGGAGTGTGTTGTTATTTGTTGTAATGTTGTTGGGACAA 2337
Db 3012 ACACAGTGTGTTGTTGGAGTGTGTTGTTATTTCTTTTGAATGTTGTTGGGACAA 3071
QY 2338 CCTCTCTTCTTGGCACAACCCCAATTTAGAAACACAAATGAAAGTGTATCATCTGCAAACT 2397
Db 3072 CCTCTCTTCTTGGCACAACCCCAATTTAGAAACACAAATGAAAGTGTATCACTGGAACA 3131

; COUNTRY: USA

1463 CGCATCAAGAGTTACTCCCTTATGCTTCAAAATCTTCATGGAGCAACACGTTGGAGAAAT 1524

QY 1444 GTCCTGAAGTCTCATCAGCAGCGCTGCTCATCGAAGAGCAGCTAGAAATGAATGATG 1503
 Db 1525 GTCTCAAAACCTACCAGCAGAGAGGTGAGCGGAGGCTACAGGTGAGCAGGAAATGGGC 1584
 QY 1504 CGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGAG 1563
 Db 1585 AAAGCTGGGCTCTGTGAGCGGAGCAGGAGCAGATGAGGAGATCTCTACAGAGAGGAG 1644
 QY 1564 TCTAATATATATTCGTCTTAAAGAGGCTTAAATGGACAAGTCTATGTTTGTAAAGATAAAG 1623
 Db 1645 TCTAATCTACAACCGGCTGAAGAGGCGCAAGATGGACAAGTCCATGTTGTGAATCAAG 1704
 QY 1624 ACATTAGGATAGAGCGTTTGTGAGTCTGTCTAGCAGAAAGTCTGATACATTAAGCT 1683
 Db 1705 ACTTAGGCTAGGCTTGGGAGTGTGCTGCTGTGTAAGCTGGACATCAAGCT 1764
 QY 1684 TTGTATGCAACAAAGACTCTTCCGAAGAAAGAGGTTCTGCTCGGAATCAGGTGGCTCAT 1743
 Db 1765 CTGTAGCCATGAAGACTCTCAGAGAGAGGATGCTGAACCGGATCAAGTGGGCCAT 1824
 QY 1744 GTCAAGCGGAGAGGATATCTCTAGCAGAGCGGACAAATGAGTGGTGGTCCGCTGTAC 1803
 Db 1825 GTCAAGGCTGAGAGGACATCTGCTGAGAGCAGACAAATCAGTGGTGGTCAAACTCTAC 1884
 QY 1804 TACTCTTCCAGCAGAGGACACTTGTACTTTGTGATGAGCTACATCTCTGGGGGGAT 1863
 Db 1885 TACTCTCTCCAGCAGAGGACAGCTGTACTTTGTGATGAGCTACATCACCAGCGGGAT 1944
 QY 1864 ATGATGAGCTATTAATAGAATGGGCTCTTCCGTAAGTCTGCAAAATCTGCACCAATCTACATA 1923
 Db 1945 ATGATGAGCTGTGATCAGGATGGAGGCTCTCCCTGAGCAGCTGGCCCTCTTACAT 2004
 QY 1924 GCAGAACTTACCTGTGCGAGTGAAGTGTTCATAAAATGGGTTTTATTCATAGAGATAT 1983
 Db 2005 GCAGAGTTGACCTGGCCATTGAAAGTGTCCACAAGATGGCTTTATCCACCGGACATC 2064
 QY 1984 AAACCTGATTAACATTTTGTATGACCGTGTGAGCCATATTAATGAGCTGACTTTGGCTG 2043
 Db 2065 AAGCCTGACAACTACTCTACGACCTGGATGTCATATTAAGCTGACAGATTTTGGCCTC 2124
 QY 2044 TGCACTGGCTTCAGATGGACATGACTCCCAAGTACTACAGAGTGGGATCACCCACGG 2103
 Db 2125 TGCACTGGATTCAGTGGAGTCAATTCACATTCACAGAAAGGAAACCATGAGA 2184
 QY 2104 CAAGATAGCATGATTCAGTACAGATGGGAGATCTTCCAAATTCCTGGTGGGGAC 2163
 Db 2185 CAGGACAGCATGAGCGCGGTGACCTCTGGGACGATGTTTCCAACTGTCTGCTGGAGAC 2244
 QY 2164 AGACTGAGCCATGGAGCGGAGAGCTGCTCGCCAGCAGCAGATGCTAGCCCATCT 2223
 Db 2245 AGTTAAAGACCTGGAGCAGAGGGCGCAGAGCAGCAGAGGTCCTGGCAGCATCT 2304
 QY 2224 CTGGTTGGGACTCCCAATTAATTTGACCTGAAGTGTCTACTGGAACAGGATATACAG 2283
 Db 2305 CTGTTGGGACACAAATTAATTCATCTCGGAGGCTGCTCTCCGAAAGGATACAGCAG 2364
 QY 2284 CTGTTGACTGGTGGAGTGTGTTATCTTTTGTGAATGTTGGTGGGACAACTCTCT 2343
 Db 2365 CTCTGTGACTGGTGGAGCGTGTGATCTCTTTGAGATGCTGTTGGGAGCGCGCT 2424
 QY 2344 TTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAAACTTCTTA 2403
 Db 2425 TTCTTGGCCCCCCCCACAGAGAGCGCAGCTGAAGGTGATCAACTGGGAGCAGCGCTG 2484
 QY 2404 CACATCCCTCTCAAGCTAAGCTGAGTCTCTGAAGCTCTGACCTCATTAATCAAACTGTGT 2463
 Db 2485 CATATCCCTACGAGGTGAGGCTCAGCGCTGAGCGCCGAGACCTCATCAGAAAGCTGTGC 2544
 QY 2464 CGAGGACCAAGAGACCGCTCGGCAAGAGCGGTGCTGCTGATGAGATAAAGGCTCATCCATTT 2523
 Db 2545 TGGCGCGCTGACTGCGCGCTGGCAGGATGGGCGAGATGACCTCAAGGACACACCGCTTC 2604
 QY 2524 TTTAAGACCATCATTTCTCTAGTGTCTGAGACAGCAGTCTGCTTCATACATCCCTAAA 2583

Db 2605 TTCAACACCATTCAGCTTTTCCCGTGACATCCGAAAGCAGGCTGACCTAGTCCCGACC 2664
 QY 2584 ATCAGCGATCCACAGATACATCAATTTTCAGCCCTGTTGATCTCTGATAAATTTGGAGC 2643
 Db 2665 ATCAGCCACCCATGGACACCTCCAAATTTTGACCGGTGATGAAGAAAGCCCTGGCAC 2724
 QY 2644 GATGGAGCAGGAGGAGGAAATATCAGTGACACTCTGAGCGGATGGTATATAAATGGGAG 2703
 Db 2725 GAGCGCAGCGGAGAG---AGCGCAAGGCTGGGACACGCTGGCCTCCCGCAGCAGCAAG 2781
 QY 2704 CACCCGAGCAGCCTTCTATGATGATCAGCTTTCAGGAGGTTTTTGTATGACAATGGCTAC 2763
 Db 2782 CATCCAGAGCAGCGCTTCTATGATGATCAGCTTCCGAGGTTCTCGATGACAAAGGCTAT 2841
 QY 2764 CCATATAATATCCAAAGCCTATTGAGTATGAATACAT 2802
 Db 2842 CCCTCCGTCGCGGAGGCTCAGAGCCCGCAGAGAGT 2880

RESULT 4
 US-09-509-902A-15
 ; Sequence 15, Application US/09509902A
 ; Patent No. 6387676
 ; GENERAL INFORMATION:
 ; APPLICANT: Virco, Duke
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Anderson, Dirk M.
 ; APPLICANT: Marken, John S.
 ; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
 ; FILE REFERENCE: 2877-US
 ; CURRENT APPLICATION NUMBER: US/09/509,902A
 ; CURRENT FILING DATE: 1999-08-03
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 1961
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-509-902A-15

Query Match 24.5%; Score 785.6; DB 4; Length 1961;
 Best Local Similarity 68.6%; Pred. No. 2.1e-204;
 Matches 1098; Conservative 0; Mismatches 499; Indels 3; Gaps 1;

QY 1220 CAGTAAGTAAAGCTGCAAGATGAACAGCTAGCTTACCCAGGAGAGATGATAGTGA 1279
 Db 115 CAGGATGAGCAGAGGCTCCGTGCGGGCCCAACGAGCCCGAGGGCGGACAGAGCC 174
 QY 1280 AGAGTGGGACAGTGTGACTCTGGGGATTAAGAAAGAAACAGATTACACTTTCACCTA 1339
 Db 175 GCAAAAGCGCAAGGGGCAAAAGGAAAGGATAAAAGCAGATTTCAGACCTCTCCCG 234
 QY 1340 TCACTGTCGAAACAAAGAAAGATCAGACAGAGAGTCTCGATTTCAGAGTTACT 1399
 Db 235 TTCCCTCCGCAAAACAGCAGACGAGAGAGAGAGAGTCCAGCATCAAGAGCTACT 294
 QY 1400 CCCACAGGCGCTTTAAGTCTTCTCATGAGCAGCAGCTAGAGAACGCTCCTCAAGTCTCATC 1459
 Db 295 CGCATACGCTTTAAGTCTTCTCATGAGCAGCAGCTGGAGATGTCATCAAAACCTACC 354
 QY 1460 AGCAGGCTGTGATCGGAGAGCAGCAGCTAGAAAATGAATGATGCGGGTGGATATCTC 1519
 Db 355 AGCAGAGGTTAAACCGGAGGCTGAGCTGGAGCAAGAAATGGCCAAAGCTCTGTG 414
 QY 1520 AAGATGCCAGGATCAAAATGAGAAAGATGCTTTCGCAAGAAAGAGTCTTAACATATATTCGTC 1579
 Db 415 AAGTGGAGGAGCAGAGTCCGGAAGTCTCTACCAAGAAAGTCTTAATACACAGGT 474
 QY 1580 TTAAGAGGCTAAATGAGCAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAG 1639
 Db 475 TAAAGAGGCGCAAGTGGACAGTCTATGTTTGTCAAGATCAAAACCTCGGGATCGGTG 534

QY 1820 AGGACAACTGTACTTTGTGATGGACTACATTCCTGGGGGGGATATGATGACCTATTAA 1879
Db 717 AAGACAGCTGTACTTTGTGATGGACTACATTCCTGGGGGGGATATGATGACCTATTAA 776
QY 1880 TTAGAATGGGATCTTCTCTGAAATCTGGACGATTTCTATAGCAGAACTTACCTGTG 1939
Db 777 TCCGATGAGGCTTCTCTGAGCAGCTGGCCGGTCTATAGCAGAGCTGACTTTGG 836
QY 1940 CAGTTGAAGTGTTCATAAATGGGTTTTATTCATAGAGATATTAACCTGATAACATTT 1999
Db 837 CCATTGAGAGTGTCCACAGATGGCTTCATCCAGGAGATCAAGCTGTATACATTT 896
QY 2000 TCATTGACCGTATGGCCATATTAATTTGACTGATTTGGCTTGTGACCTGGCTTCAGAT 2059
Db 897 TGATAGATCTGGTGGTTCACATTAATCAAGATTTCCGCTCTGACCTGGTTCAGT 956
QY 2060 GGACATGACTTCCAAAGTACTACAGAGTGGGATCAACCCAGGCAAGATAGCATGGATT 2119
Db 957 GGACTCAAAATTCACAAATTTACAGAAAGGAGGACCATGTACAGAGCAGCATGGAGC 1016
QY 2120 TCAGTAACGAATGGGAGATCTTCCAAATTTGCTGGTGGGGAGAGACTGAAGCCACTGG 2179
Db 1017 CCAGGACCTCTGGGATGATGTCTTAACCTGGTGGGGAGAGCTGAAGACCCCTAG 1076
QY 2180 AGCGGAGAGTGTCTGCGCAGCAGCAGGATGCTAGCCCATTTCTGGTGGGACTCCCA 2239
Db 1077 AGCAGAGGGCGGAAGCAGCAGCAGGATGCTGGCAGCATTCACCTGGTGGGACTCCAA 1136
QY 2240 ATTATATTGACCTGAAGTGTCTGCGAAGAGGATATACAGCTGTGTGAGTGGTGA 2299
Db 1137 ACTACATCGGACCGAGGTCTCTCGGCAAGGGTACACTCACTCTGACTGTGGA 1196
QY 2300 GTGTGTGTATTTCTTTGGAATGTGTGGGAGCAACTCTTCTTGGCAGCAAAACC 2359
Db 1197 GTGTGGAGTGTCTTTCAGATGCTGTGGGGAGCGCCCTTTTGGCAGCTACTC 1256
QY 2360 CATTAGAACACAAATGAAGTTATCATCTGGCAAACTTCTTACACATCCCTCTCAAG 2419
Db 1257 CCACAGAAACCCAGCTGAAGGTGATCACTGGGAGAACAGCTCCACATTCAGCCCAAG 1316
QY 2420 CTAAGCTGAGTCTGAAGCTCTGACCTTATTAACACTGTGTGCGAGGACCAAGACC 2479
Db 1317 TGAAGCTGAGCTGAGGCGGAGGACCTCATCAACAGCTGTGTCTCCGACAGACC 1376
QY 2480 GCCTCGGCAAGACGGTGTGTAGATGAAGGCTCATCCATTTTAAAGACCATCGATT 2539
Db 1377 GCCTGGGCGGAATGGGCGGATGACCTGAAGGCGCCACCCCTTCTTACGCCCATTTGACT 1436
QY 2540 TCTTAGTGTATGACAGCAGAGTCTGCTTATACATCCCTTAAATCAGCATCCACAG 2599
Db 1437 TCTCCAGTGCATCCGGAAGCATCCAGCCCTTAGGTTCCACCATCAGCCACCCCATGG 1496
QY 2600 A 2600
Db 1497 A 1497

RESULT 6
US-09-442-100-1
; Sequence 1, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weli
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
NAME/KEY: CDS
LOCATION: 1103..4402
US-09-442-100-1

Query Match 16.2%; Score 521.2; DB 4; Length 5720;
Best Local Similarity 65.4%; Pred. No. 5.5e-132;
Matches 852; Conservative 0; Mismatches 423; Indels 27; Gaps 5;

QY 1354 AACAGAAAGATCAAGAACGAGAGAGTCTCGGATTCAGAGTTACTCCACAGGCGCTTT 1413
Db 2984 AAGGAGAGGAGGAGGAGGAGGAGGAGTTCGCGCATCAGCAGTACTCGCGCAAGCGCTC 3043
QY 1414 AAGTCTTCTGAGGAGCAGCAGTGTAGAGAGCTCTGAAGTCTCATCAGCAGGCTGTGAT 1473
Db 3044 AAGTCTTCTGAGGAGCAGCAGTGTAGAGAGCTGTATCGGATCGTATCGGCGGACGAT 3103
QY 1474 CGGAAAGCAGCTAGAAATGAAATGATGCGGGTGGATTATCTCAAGATGCCAGGAT 1533
Db 3104 CGGAAAGTCACTGGAGAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCAAATC 3163
QY 1534 CAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAATATATTCGTTTAAAGGGCTAAA 1593
Db 3164 GAGATCAGGAAATCTGAAACCAAGAGGAGAGCAACTACATTCGATTGAAGCGGCGCAAG 3223
QY 1594 ATGGCAAGTCTATCTTTGTAAGATAAAGACATTAGGAATAGGAGCGTTTGGTGAAGTC 1653
Db 3224 ATGGCAAGAGCATGTTCTGCTCAAACTGAAGCCCATTTGAGTGGTGTGATTTGGCAGGTA 3283
QY 1654 TGCTCTAGCAAGAAAGTCGATA---CTAAAGCTTTGTATGCAACAAAGACTCTTTTCGAAG 1710
Db 3284 AGCTGTGGTGAACAAATCGATACCTCGAAACCAATTTGTATGCGATGAAACCCCTCGGAAA 3343
QY 1711 AAGACGTTCTGCTCCGAAATCAGGTGGCTCATGTGAAGCGGAGAGGATATCTTAGCA 1770
Db 3344 GCGGACGTTCTCAAGCGGAATCAGGTGGCAGACGCTGAAGCGGAGAGGATATCTTCGCG 3403
QY 1771 GAAGCCGACATGAGTGGTGGTCCGCTGTACTACTCTTTTCCAGGACAAAGCAACTTG 1830
Db 3404 GAAGCCGACAAATGAGTGGTGGTGAAGTTGTACTACTACAGCTTCCAGGACAAAGGATAATCTG 3463
QY 1831 TACTTTGTGAGTACTACATTCCTGGGGGGGATATGATGAGCCTATTATTAAGATGGGC 1890

RESULT 8
US-09-328-111-66
; Sequence 66, Application US/09328111
; PATENT NO. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.

97	ATTAAACCTGATAATATTTTGATTGATCGTGATGTCATATT	db
1981	ATTAAACCTGATACATTATTTTGATGACCGTAGTGCCCAAT	db
2041	TTGTGCACCTGGCTTCAGATGGACACATGACTCCAAGT	db
37	CTCTGCACCTGGCTTCAGATGGACACAGATTCCTAAGT	db

APPLICANT: Bushnell, Steven E.
 APPLICANT: Carroll III, Eddie
 APPLICANT: Catino, Theodore J.
 APPLICANT: Derti, Adnan
 APPLICANT: Ford, Donna M.
 APPLICANT: Lewis, Marcia E.
 APPLICANT: Monahan, John E.
 APPLICANT: Schlegel, Robert
 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 FILE REFERENCE: CDD-257 (US)
 CURRENT APPLICATION NUMBER: US/09/328,111
 EARLIER FILING DATE: 1999-06-08
 EARLIER APPLICATION NUMBER: US 60/088,801
 EARLIER FILING DATE: 1998-06-10
 NUMBER OF SEQ ID NOS: 850
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 66
 LENGTH: 678
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc.feature
 LOCATION: (1)...(678)
 OTHER INFORMATION: n = A,T,C or G
 US-09-328-111-66

Query Match 11.7%; Score 376.6; DB 4; Length 678;
 Best Local Similarity 82.4%; Pred. No. 4.8e-93;
 Matches 500; Conservative 0; Mismatches 97; Indels 10; Gaps 6;

QY 1447 CTCAAGTCTCATCAGCAGCGTCTGCATCGGAAGAAGCAGCTAGAAAATGAATGATCGG 1506
 DB 2 CTCAAATCTCATCAGCAGCGTCTCATCGTAATAAACAATTAGCAATGAATGATCGG 61

QY 1507 GTTGGATTATCTCAAGATCCCGAGATCAATGAGAAAGATGCTTGGCCAGAAAGCTCT 1566
 DB 62 GTTGGATTATCTCAAGATCCCGAGATCAATGAGAAAGATGCTTGGCCAAAGAAATCT 121

QY 1567 AACTATATTCGTCTAAAGGCTAAATGACAAAGTCTATGTTTGTAAAGATAAGACA 1626
 DB 122 ATTACATCCGCTCTAAAGGCTAAATGACAAAGTCTATGTTTGTAAAGATAAGACA 181

QY 1627 TTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATCTAAAGCTTG 1686
 DB 182 CTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTCTAGTCTAGGCTTTG 241

QY 1687 TATGCAACAAAGACTCTTCGAAAGAACAGCTTCTGCTCCGAAATCAGTGGCTCATGTG 1746
 DB 242 TATGCAACAAAGACTCTTCGAAAGAACAGATGTTCTTCTCGAAATCAAGTCGCTCATGT 301

QY 1747 AAGCGGAGAGGATATCTTACGAGAGCCGACAAATGAGTGGTGGTCCGCTCTACTAC 1806
 DB 302 AAGCGTGAGAGATATCTTCTGAGCTGACAAATGAGTGGTGGTGGTGGTGGTGGTGGT 361

QY 1807 TCTTCCAGGACAGGACAACTTCTTGTGATGACTACATCTCTGGGGGGGATATG 1866
 DB 362 TCATTCCAGGACAGGACAACTTCTTGTGATGACTACATCTCTGGGGGGGATATG 419

QY 1867 ATGAGCTTATTAATAGATGGGATC-TTCTCTGAAATCTGGCAGGATCTACAT--A 1923
 DB 420 AAGAGCCATTAATTAATAGATGGGATC-TTCTCTGAAATCTGGCAGGATCTACAT--A 479

QY 1924 GCAGAACTTACCTGTG-CAGTTGAAAGTGTTCAT--AAATGGGTTTATTCATAGAGAT 1980
 DB 480 CCAGAACTTACCTGTG-CAGTTGAAAGTGTTCAT--AAATGGGTTTATTCATAGAGAT 539

QY 1981 ATTAACCTGATACATTTTGGATG--ACCGTATGCGCCATTAATTAATGACTGCTTTG 2038
 DB 540 TTTTAACTGATATTTTGGATG--ACCGTATGCGCCATTAATTAATGACTGCTTTG 2099

QY 2039 GCTTGTG 2045
 |||

DB 600 GCCTTNG 606

RESULT 9

US-08-878-989-11
 ; Sequence 11, Application US/08878989
 ; Patent No. 5885803
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl G.
 ; APPLICANT: Lal, Preeti
 ; APPLICANT: Goli, Surya K.
 ; APPLICANT: Shah, Purvi
 ; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/878,989
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0321 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; TELEX:
 ; INFORMATION FOR SEQ ID NO: 11:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1935 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: SYNORAT04
 ; CLONE: 705365
 ; US-08-878-989-11

Query Match 7.7%; Score 248.4; DB 2; Length 1935;
 Best Local Similarity 53.2%; Pred. No. 8.3e-58;
 Matches 599; Conservative 0; Mismatches 491; Indels 36; Gaps 2;

QY 1414 AAGTCTTCTCATGAGCAGCAGCTAGAGAACTCTGAGTCTCATCAGCAGCGTCTGCAT 1473
 DB 122 AAGTCTTCTCATGAGCAGCAGCTAGAGAACTCTGAGTCTCATCAGCAGCGTCTGCAT 181

QY 1474 CGAAGAAAGCAGCTAGAAAATGAATGATCGGGTGTGATTTATCTCAAGATGCCAGAT 1533
 DB 182 AGGAGAGAAATAGAGTGGCCATGGAGAGAGAGATAGCAGATGAAGAGAAAAG 241

QY 1534 CAATGAGAAAGATGCTTTGCCAGAAAGAGTCTTAACATATATTCCTCTTAAAGGGCTAAA 1593
 DB 242 TTACGTGATCAACACACGCTCGCAAGAAAGAGAGTCTTTACGGCTCAAAAGGACCAGA 301

QY 1594 ATGACAAGTCTATGTTTGTAAAGATAAGACATTAGGAATAGGAGCGTTTGTGAGTC 1653


```

Db 302 CTGCTGGTGGTACTTTGAGTCTCTGAAAGTTATAGAAAGAGGAGCTTTTGGAGAGTG 361
Qy 1654 TGTCTAGCAAGAAAGTCTGATCTAAAGCTTTGTATGCAACAAAGACTCTTCTGAAAGAAA 1713
Db 362 CGTGTGGTCCACAAAAGATACAGGCCATATCTATGCAATGAAGATATTGGAAGTCT 421
Qy 1714 GACGTTCTGCTCCGAAATCAGTGGCTCATGTGAAGCGGAGGAGATATCTTACAGAAA 1773
Db 422 GATATGCTTTGAAAAGAGAGAGTGGCCCATATCCAGCAGAAAGAGATATTGTTGTAGAA 481
Qy 1774 GCCGACATGAGTGGTGGTCCGCTGTACTACTCTTTCCAGGACAAAGCAACTTGTAC 1833
Db 482 GCAGATGCTGGTGGTGGTGAAGATGTTTACAGTTTTCAGGATTAAGAGGAATCTTTAT 541
Qy 1834 TTTGTGATGGTACTACATCTCTGGGGGGATGATGAGCTGATTAATTAAGATGGGCATC 1893
Db 542 CTAATCATGGAATTTCTCCCTGGAGTGCATGATGATGATGATGATGATGATGATGATGAT 601
Qy 1894 TTTCTGAAATCTGGCAGCATTTCTACATAGCAGAACTTACCTGTGCAAGTTGAAAGTGT 1953
Db 602 TTGACAGAAGAGGAACACAGTCTTACATTTTCAGAGACTGTTCTGGCAATAGATCGCATC 661
Qy 1954 CATAAATGGTGTATTCATAGAGATATTAACCTGTATACATTTTGATGACCGTGTAT 2013
Db 662 CACCAGTTGGTGTATTCATCCATCGGATATTAAGCCAGACAACTTTTATGATGCCAAG 721
Qy 2014 GGCATATTAATGACGTACTTTGGCTGTGCACTGCTTTCAGATGACACATGACATGCC 2073
Db 722 GGTATGTAATTAATCATCTGATTTGGTTCATGTCAGGATTAAGAAAGCTCACAGGACT 781
Qy 2074 AAGTACTACAGAGTGGGATCACCACCGCAAGATAGCATGATTTTCAGTAACCAATGG 2133
Db 782 GAATTTTAT-----AGAAATCTCACACACACCAACCAAGT 818
Qy 2134 GGAGATCTTCCAATTTGCTGGTGGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCT 2193
Db 819 ACTTCTCATTTTCAGAACATGAATCAAAAGAGAAAGCAAACTTTGGAAGAAGA----- 872
Qy 2194 CGCCAGCACACGAGTCTAGCCCATCTCTGGTGGGACTCCCAATTAATGACACT 2253
Db 873 -----ACAGAGACAACTGGCATATTCACAGTTGGGACACAGATTAATGCTCCA 925
Qy 2254 GAAGTGTCTACGCAACAGGATATACACAGCTGTGCTGAGTGGTGGTGGTGGTGGTGGTGGT 2313
Db 926 GAAGTATTCATGACAGCTGGTTTACAACTAATTTGTGACTGGTGGTGGTGGTGGTGGTGGT 985
Qy 2314 CTTTGTGAATGTTGGTGGGACACCTCTCTTTTGGCACAACCCCATTAAGAAACAA 2373
Db 986 ATGATGAATGCTAATAGGATATCCACCTTTCTGCTCTGAAACACCTCAAGAAACATAC 1045
Qy 2374 ATGAAGGTATCATCTGGCAACTTCTTACACATCCCTCCTCAGCTAAGCTGAGTCTCT 2433
Db 1046 AGAAAGTGTGACTGGAAGAACTCTGGTATTTCTCCAGAGGTACCTATATCTGAG 1105
Qy 2434 GAAGCCTCTGACCTCATTTATCAAACTGTGTGAGGACAGAGAGCGCTCGGCAAGAAC 2493
Db 1106 AAAGCAAGGACTTAATTTCTCAGATTTTGTATGATTTCTGAAACAGAAATTTGGAATAGT 1165
Qy 2494 GGTCTGATGATGAATAAGGCTCATCCATTTTAAAGCCATCGATT 2539
Db 1166 GGAGTAGAAGAAATAAAGGTGTCATCCCTTTTGAAGGTGTCGACT 1211

```

RESULT 10

US-09-272-796-11
 ; Sequence 11, Application US/0927296
 ; Patent No. 6207148
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl G.

APPLICANT: Lal, Preeti
 APPLICANT: Goli, Surya K.
 APPLICANT: Shah, Purvi
 TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
 TITLE OF INVENTION: KINASES
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Incyte Pharmaceuticals, Inc.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/272,796
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/878,989
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0321 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 TELEX:
 INFORMATION FOR SEQ ID NO: 11:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1935 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: SYNORAT04
 CLONE: 705365
 US-09-272-796-11

Query Match 7.7%; Score 248.4; DB 4; Length 1935;
 Best Local Similarity 53.2%; Pred. No. 8.3e-58;
 Matches 599; Conservative 0; Mismatches 491; Indels 36; Gaps 2;

```

Qy 1414 AAGTTCTTCATGGAGCAGCAGCTAGAGAACGCTCTCTGAAAGTCTCTCATCAGCAGCGTCTGCAT 1473
Db 122 AAGCTCACATTGGAGAAATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAAACC 181
Qy 1474 CGAAGAGCAGCTAGAAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGAT 1533
Db 182 AGGCAAGAGAAATTAGAAGTGGCCATGGGAAGAAGGATTAGCAGATTGAAGAGAAAAG 241
Qy 1534 CAATGAGAAAGATGCTTTGGCAGAAAGAGCTACTATATTCTGCTTTTAAAGGGCTAAA 1593
Db 242 TTACGTCGATCACACACGCTCGCAAGAACAGAGTCTTACGGCTCAAAAGGACCAGA 301
Qy 1594 ATGCAAGAGTCTATGTTGTTAAAGATAAAGACATTTAGGAATAGGAGCGCTTTGGTGAAGTC 1653
Db 302 CTTGGCTTGGATGACTTTTGGAGTCTCTGAAAGTTATAGGAAGAGGAGCTTTTGGAGAGTG 361
Qy 1654 TGTCTAGCAAGAAAGTGCATATCACTAAAGCTTTGTTGATGCAACAAAGACTCTTCTGAAAGAAA 1713
Db 362 CGGTTGGTCCCAAAAAAGATACAGGCCATATCTATGCAATGAAGATATTGGAAGTCT 421
Qy 1714 GACGTTCTGCTCCGAAATCAGTGGCTCATCTGTAAGCGGAGAGGAGATATCTTACAGAAA 1773
Db 422 GATATGTTGAAAAGAGCAGGTTGGCCCATATCCAGCAGAAAGAGATATTGTTGTAGAA 481
Qy 1774 GCCGACATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1833

```

QY	2138	ATCCTTCCAAATTGTCGGTGTTGGGGACAGACTGAAGCCACTCGACGCGAGAGAGCTGCTCGCC	2149
Db	1379	-----ATGAATTCCAAAGAAAGCAAAACCTGGAAAAG	1411
QY	2198	AGCACAGCATGTCTAGGCCATTTCTGGTGGGACTCCCAATTAATTGCACCTGAAG	2255
Db	1414	A-AATAGAGCTCAGCTTAGCCTTCTCCACAGTAGGCACCTCCTGACTACATTGCTCCTGAGG	1476
QY	2258	TGCTACTTGGGAACAGGATATACACAGCTGTGTGACTGTGGAGTGTGGTGTTATCTTTT	2315
Db	1473	TGTTTATGCAGACCGGGTACAACAAGCTCTGTGATTGTGGTTCGGCTGGGGGTGATCATGT	1530
QY	2318	GTCNAATGTGGTGGACAACTCTCTTCTGGCACAAACCCTATTAGAACAACAATAA	2375
Db	1533	ATGAGATGTCATCGCGCTACCCACCCTTCTGTCTTGAGACCCCTCAACAGACATATAA	1590
QY	2378	AGGTTATCATCTGGCAAACTTCTTACACATCCCTCCTCAAGCTAAGCTGAGTCTCTGAAG	2435
Nb	1593	AGGTGATGAACCTGGAAGAAACCTTTCGACTTTTCTCCGAAAGTCCCATTCTTGAGAAG	1650

QY 2438 CCTCTGACCTCATTATCAAACTGTCTGAGGACCAAGACCGCTCGCAAGAACGGTG 2497
 Db 1653 CCAAGGATCTAATTTTGGTCTCTGCTGAATGGAACATAGATTGGAGCTCTGGAG 1712
 QY 2498 CTGATGACATAAAGCTCATCCATTTTAAAGACCATCGATTTCTCTAGTGTGAGAC 2557
 Db 1713 TTGAGGAATAAAGAGTAACCTCTTTTGAAGGCGTTGA---CTGGAACATATCAGAG 1769
 QY 2558 AGCAGCTGCTTCATACATCCCTAAATCACGCATCCCAACAGACATACCAATTCGA 2615
 Db 1770 AGAGACCTGCTGCATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTCGA 1827

RESULT 12

US-09-338-132-6

; Sequence 6, Application US/09338132

; Patent No. 6040164

; GENERAL INFORMATION:

; APPLICANT: Hemmings, Brian A.

; APPLICANT: Millward, Thomas A.

; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases

; FILE REFERENCE: 14-20265/A/PCT

; CURRENT APPLICATION NUMBER: US/09/338,132

; CURRENT FILING DATE: 1999-06-22

; EARLIER APPLICATION NUMBER: 08/860,150

; EARLIER FILING DATE: 1997-06-19

; EARLIER APPLICATION NUMBER: PCT/EP95/05052

; EARLIER FILING DATE: 1995-12-20

; EARLIER APPLICATION NUMBER: 94810746.1

; EARLIER FILING DATE: 1994-12-22

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 3018

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (596)..(1990)

US-09-338-132-6

Query Match 7.6%; Score 244.4; DB 3; Length 3018;
 Best Local Similarity 53.4%; Pred. No. 1.3e-56;
 Matches 608; Conservative 0; Mismatches 491; Indels 39; Gaps 3;

QY 1478 AGAACAGCTAGAAATGAATGATCGCGGTGGATTATCTCAAGATGCCAGGATCAA 1537
 Db 729 AAAGAAGTTAGAAAAGTGTGGAAGAGAGGCGCTAAAAGATGAGGAGAAACACTCC 788
 QY 1538 TGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTTTAAAAGGGCTAAAATGG 1597
 Db 789 GGAGATCAGCATCTCGGAAGGAAACAGAGATTTCTGTTTGAAGAGAAACAGACTTG 848
 QY 1598 ACAAGCTATGTTTGAAGATAAAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTC 1657
 Db 849 GATTGGAAGATTTTTCAGTCTTAAAAGTAATAGGACAGAGGACATTTGTTGAGGTACGGC 908
 QY 1658 TAGCAAGAAAGCTCATACTAAAGCTTTGTATGCAACAAAGACTCTTCAAGAAAGAGC 1717
 Db 909 TTGTTTCAAGAAAGATACGGGACATGTGTATGCAATGAAATATCTCCGTAAAGCAGATA 968
 QY 1718 TCTCTCCGAAATCAGTGGCTCATGTGAAAGCGGAGGATATCTTAGCAGAGCGG 1777
 Db 969 TGCITGAAAACAGACAGGTTGCCACATTCGTCGCGAGCGTGACATTTCTAGTGGAGGAG 1028
 QY 1778 ACAATGAGTGGTGGTCCGCTGTACTACTCTTTCCAGGACAAAGCAACTTTGTACTTTG 1837
 Db 1029 ACAGTTTGTGGTCTGTAAGATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTACCTAA 1088
 QY 1838 TGATGACATACATTCCTGGGGGGATGATGAGGCTTATTAATAGATGGGCATCTTTC 1897
 Db 1089 TCATGGAGTTCTGCTGGAGGGACATGATGACCTTGTGTATGATGAAAAGACACTCTGA 1148

QY 1898 CTGAAAATCTGGCAGCAATTTACATAGCAGAACTTTACCTGTGCACTGTGAAAGTGTTCATA 1957
 Db 1149 CAGAGAGAGAGACTCAGTTTATATAGCAGAACTAGTATTAGCCATAGACTTATTCACC 1208
 QY 1958 AAATGGGTTTATTCATAGAGATATTAACCTGATAAATTTTGTAGTACCGGTGATGCC 2017
 Db 1209 AACTTGGGATTCATCCACAGAGACATCAACACAGAACTCTCTTTTGGACAGCAAGGCC 1268
 QY 2018 ATATTAAATTTGCTACCTTCTGCTTGTGCACTGCTTGTGAGTGGACAGATGATCCAAAT 2077
 Db 1269 ATGTGAAATTTCTGACCTTTTGTGCTTGTGACAGGAGTGAAGAAACAGATAGGACAGAT 1328
 QY 2078 ACTACAGAGTGGGATCACCCACGCAAGATAGATGATTTTCACTAACCAATGGGAG 2137
 Db 1329 TTTATAGGANTCTGAACACAGCTCCCCAGTGAATTTTCCAGAAC----- 1378
 QY 2138 ATCTTCCAAATTTGCTGGTGGGACAGACTGAAGCCACTTGGAGCGGAGAGCTGCTCGCC 2197
 Db 1379 -----ATGAATTTCCAAAAGGAAAGCAGAAACCTGGAAAAG 1413
 QY 2198 AGCACCAGGATGTCTAGCCCAATTTCTGTTGGGACTCCCAATTTATTTGACCTGAAG 2257
 Db 1414 A-AATAGACGTGAGTACCTTTCTCCAGTAGGCACTCTGACTACATTTGCTCTGAGG 1472
 QY 2258 TGCTACTCGAACAGGATATACACAGCTGTGTGACTGTGGAGTGTGTGTTTATTTCTTT 2317
 Db 1473 TGTTCATGACAGCCGGTACACAGCTGTGTGATTGGTGGTCTGCTGGGTGATCATCT 1532
 QY 2318 GTGAAATTTGGTGGGCAACCTCTTTCTTGGGCAACAAACCCATTTAGAAACAAATGA 2377
 Db 1533 ATGAGATGCTATCGGCTACCCACCTTTCTGTTCTGAGACCCCTCAAGAGACATATAAGA 1592
 QY 2378 AGGTTATCATCTGGCAAACTTCTCTACACATCCCTCAAGCTAAGCTGAGTCTCTGAAG 2437
 Db 1593 AGGTGATGAATGGAAGAAACTTTGACTTTTCTCCAGAGATTTCCCTCTGAGAAAG 1652
 QY 2438 CCTCTGACCTCATTATCAAACTGTGTGAGGACCAAGAGACCGCTCGCAAGAAACGGTG 2497
 Db 1653 CCAAGGATCTAATTTTGGTGTCTGCTGTAATGGAAACATAGAAATGGAGCTCTGGAG 1712
 QY 2498 CTGATGAGATAAAGCTCATCCATTTTAAAGACCATCGATTTCTCTAGTGTGAGAC 2557
 Db 1713 TTGAGGAATAAAGAGTAACCTCTTTTGTGAGGCGTTGA---CTGGAACATATCAGAG 1769
 QY 2558 AGCAGTCTGCTTCATACATCCCTTAAATCACGCATCCCAACAGACATACCAATTTTGA 2615
 Db 1770 AGAGACCTGCTGCAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTCGA 1827

RESULT 13

US-08-860-150-1

; Sequence 1, Application US/08860150B

; Patent No. 5981205

; GENERAL INFORMATION:

; APPLICANT: Hemmings, Brian A.

; APPLICANT: Millward, Thomas A.

; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases

; FILE REFERENCE: 4-20265/A/PCT

; CURRENT APPLICATION NUMBER: US/08/860,150B

; CURRENT FILING DATE: 1997-06-19

; EARLIER APPLICATION NUMBER: PCT/EP95/05052

; EARLIER FILING DATE: 1995-12-20

; EARLIER APPLICATION NUMBER: 94810746.1

; EARLIER FILING DATE: 1994-12-22

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2101

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (132)..(1499)

US-08-860-150-1

Query Match 6.2%; Score 198.6; DB 2; Length 2101;
Best Local Similarity 51.5%; Pred. No. 3,4e-44;
Matches 552; Conservative 0; Mismatches 489; Indels 30; Gaps 3;

QY 1414 AAGTCTTTCATGAGCAGCAGTACAGACGCTCTGAGTCTCATACAGCAGCTCTGCAT 1473
DB 213 AAGGTGAGCTTGGAGAACTACTACAGCAACCTGCTGAGCAGTATGCGGCGGAAAGCAG 272
QY 1474 CGAAGAAGCAGCTAGAAAATGAATGATGCGGTGGATTATCTCAAGATGCCAGAT 1533
DB 273 CGCTCGCAAGCTGGAGCTCAGCTGAGCAGGAGCTTGTGCGAGGCGCAGCGCAG 332
QY 1534 CAATATGAGAAAGATGCTTGGCAGAAAGTCTAACTATATTCCTCTTAAAGGGCTAAA 1593
DB 333 GAGAGGCTCTGCAGCATGCCAGAGGAGAGCGAGTATCTCCGCTGAAGCGATTGGCG 392
QY 1594 ATGACAGCTCTATGTTTGAAGATTAAGACATTAAGATAGAGCGTTTGGTGAAGTC 1653
DB 393 CTCGGTGTGGAGGACTTTGAGCGCTCAAAAGTATCGACGCGCGGCTTCGGTGAAGTG 452
QY 1654 TGTCTAGCAAGAAAGTCGATACCTAAAGCTTGTATCAACAAGACTCTTCGAAAGAA 1713
DB 453 CGTTTGTGCAAGAAAGGACACTGGACATGTGTGCGCCATGAAGGTCTGCGCAAGCG 512
QY 1714 GACGCTTCTGCCAAATCAGGTGGCTCATGTGAAAGCGGAGAG--GGATATCTAGCA 1770
DB 513 GACATGCTGGAAGAGGAGCAGGTGGCAGCAGCTAGCGCCGAGGGTCTGATGCTCGGTC 572
QY 1771 GAAGCGGACAAATGAGTGGGTGGTCCGCTCTACTCTTCCAGGACAGGACAACTTG 1830
DB 573 GAGGCGCATCATCAGTGGGTGGTGAAGTGTACTACAGTTCGAGGATCCCGTCAATTTA 632
QY 1831 TACTTCTGATGACTACATTCCTGGGCGGATATGATGAGCCTTAAATAGATGGCG 1890
DB 633 TATTGATATGAGTCTTTCGCTGGGTGGTATGATGAGCTTTTAAATGAAGAGGAC 692
QY 1891 ATCTTCTCCTGAAATCTGGCAGCACTTACATAGCAAACTTACCTGTGCAAGTTGAAAGT 1950
DB 693 ACGTATCCGAGGAGGACACACAGTTCTATATCAGTGAGAGCGCATGGCGATTCGATTCT 752
QY 1951 GTTCATAAATGGTTTTTATTCATAGAGATATTAACCTGATTAACATTTTGGATGACCGT 2010
DB 753 ATTACAAATCGGTTTATACACAGGATATCAAGCCGATAACTTGTGCTGGAGCG 812
QY 2011 GATGGCCATATTAATTTGACTGCTTGGCTTGTGCACTGGCTTCAGATGAGCAGATGAC 2070
DB 813 CGAGGCTATCGAGCTCTCGACTTCGACTGTGCACTGGCTTAAGAAAGTGCATCGA 872
QY 2071 TCCAAGTACTACAGAGTGGGATCACCCAGCGCAAGATAGCATGGATTTCAGTAACGAA 2130
DB 873 ACAGACTTTTA-----TCGGGACTTTCGAGCGCAAAACCATCCGATTTTATAGGCA 926
QY 2131 TGGGAGATCCTTCCATTTGCTGGTGGGAGCAGACTGAAGCCACTGGAGCGGAGGCT 2190
DB 927 TGGCCAGTCCGATGGACTCCAAAGCGAGCTGCGGAGTCTGGAAGCGAAATTCAGCGCGC 986
QY 2191 GCTCGCCAGCAGCAGGATGCTAGCCCATCTCTGTTGGGACTCCCAATTTATTTGCA 2250
DB 987 -----CTCGCTACAGCAGCAGCTGGGAGCGCGGACTATATTGCA 1025
QY 2251 CTGGAAGTGTACTGGCAACAGGATATACAGCTGTGACTGGTGGAGTGTGGTGT 2310
DB 1026 CCGGAAGTATTTTCAGAGACTGGTACGAGCCCGCTGCGACTGGTGGTTCCTCGGAGTC 1085
QY 2311 ATTCTTGTGAATGTTGGTGGACACACTCTTCTTGGCACAACCCCAATTAGAAACA 2370
DB 1086 ATCATGTACGAATATGATGGGCTATCTTCCATTTGCTCGGACAAATCCCGAGGACAC 1145
QY 2371 CAAATGAAGTGTATCTCTGGCAAACTTCTACATCCATCCCTCTCAAGCTAAGCTGAGT 2430
DB 1146 TACCGCAAGGTGATGAATGCGCGGAGACGCTGATATTTCCCGCAAGAGATCCCATATCG 1205

QY 2431 CCTGAAGCCTGTGACCTCAATTATCAAACTGTGTCGAGGACAGAAAGACCGC 2481
DB 1206 GAGAGCGCAAGGAGACGATCACTCACTTCTGTCGAGCGCGATCGCGC 1256

RESULT 14

US-09-338-132-1
Sequence 1, Application US/09338132
Patent No. 6040164
GENERAL INFORMATION:
APPLICANT: Hemmings, Brian A.
APPLICANT: Millward, Thomas A.
TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
FILE REFERENCE: 4-20265/A/PCT
CURRENT APPLICATION NUMBER: US/09/338,132
CURRENT FILING DATE: 1999-06-22
EARLIER APPLICATION NUMBER: 08/860,150
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05052
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: 94810746.1
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 2101
TYPE: DNA
ORGANISM: Drosophila melanogaster
FEATURE:
NAME/KEY: CDS
LOCATION: (132)..(1499)
US-09-338-132-1

Query Match 6.2%; Score 198.6; DB 3; Length 2101;
Best Local Similarity 51.5%; Pred. No. 3,4e-44;
Matches 552; Conservative 0; Mismatches 489; Indels 30; Gaps 3;

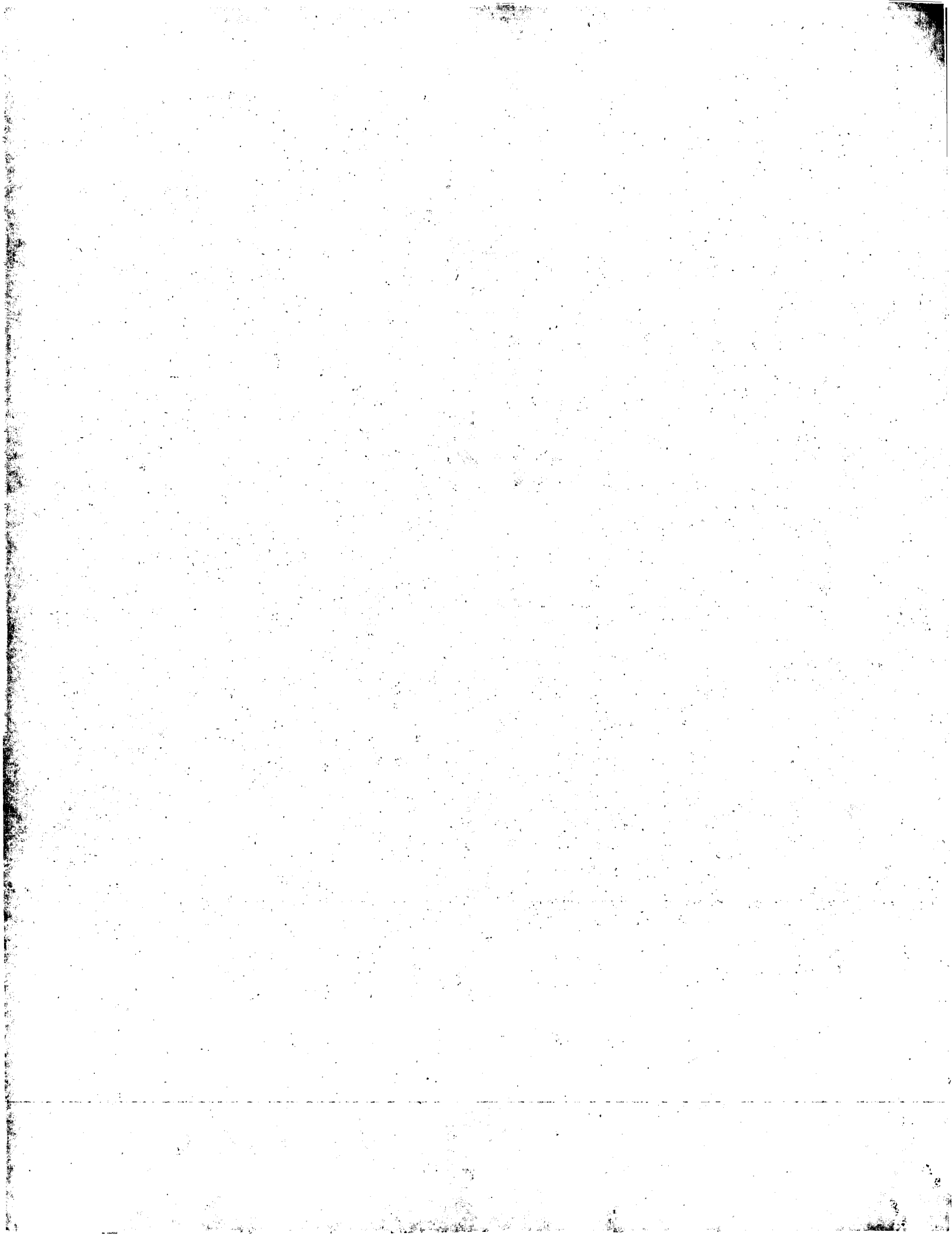
QY 1414 AAGTCTTTCATGAGCAGCAGTACAGACGCTCTGAGTCTCATACAGCAGCTCTGCAT 1473
DB 213 AAGGTGAGCTTGGAGAACTACTACAGCAACCTGCTGAGCAGTATGCGGCGGAAAGCAG 272
QY 1474 CGAAGAAGCAGCTAGAAAATGAATGATGCGGTGGATTATCTCAAGATGCCAGGAT 1533
DB 273 CGCTCGCAAGCTGGAGGCTCAGCTGAAGGAGCAGAGCTTGTGCGGCGCGCAGCGCAG 332
QY 1534 CAAATGAGAAAGTCTTTCGCCAGAAAGTCTAACTATATTCGCTTAAAGCGCTAAA 1593
DB 333 GAGAGGCTGTGAGCATGCCAGAGGAGCAGGATATCTCCGCTGAAGCGGATTGGCG 392
QY 1594 ATGACAAAGTCTATGTTTGTAAAGATAAGACATTAAGGATAGGAGCGTTTGGTGAAGTC 1653
DB 393 CTCGGTGTGGAGGACTTTGAGGCGCTCAAAAGTCAATCGACGCGCGGCTTTCGGTGAAGTG 452
QY 1654 TGTCTAGCAAGAAAGTCTGATACTAAAGCTTGTATGCAACAAGACTCTTCGAAAGAAA 1713
DB 453 CGTTTGTGCAAGAAAGGAGCAGCTGGACATGTGTGCGCCATGAAGTGTCTGCGCAAGCG 512
QY 1714 GAGTCTTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGAG--GGATATCTAGCA 1770
DB 513 GACATGCTGGAAGAGGAGCAGGTGGCAGCAGCTAGCGCCGAGGGTCTGATGCTCGGTC 572
QY 1771 GAAGCGGACAAATGAGTGGGTGGTCCGCTCTACTCTTCCAGGACAGGACAACTTG 1830
DB 573 GAGGCGCATCATCAGTGGGTGGTGAAGTGTACTACAGTTCGAGGATCCCGTCAATTTA 632
QY 1831 TACTTGTGATGACTACATTCCTGGGCGGATATGATGAGCCTTAAATAGATGGCG 1890
DB 633 TATTGATATGAGTCTTTCGCTGGGTGGTATGATGAGCTTTTAAATGAAGAGGAC 692
QY 1891 ATCTTCTCCTGAAATCTGGCAGCACTTCTACATAGCAAACTTACCTGTGCAAGTTGAAAGT 1950
DB 693 ACGTATCCGAGGAGGACACACAGTTCATATCAGTGAGAGCGCATGGCGATTCGATTCT 752

QY	1569	CTATATTTCGTCCTTAAAGGGCTTAAATGGACAGTCTATGTTTCTGAAGATAAAGACATT	1628
Db	888	GTTTCTGCGCTTCGCTGAGACACGGCTATCCCTGGAAGATTTCACACTGTGTAAGTCAT	947
QY	1629	AGGAATTAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAAGTCGATACTAAAGCTTTGTA	1688
Db	948	AGGAAGGGTGCAATTCGGTGAGGTCGCTCGTGGTCGAAGAAGAAAGATACCGGTAAATATA	1007
QY	1689	TGCAACAAAGACTCTTCGAAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCTCATGTGAA	1748
Db	1008	CGCTATGAAGACATGTTTAAAAATCAGAATGTACAAGAAGGATCAATTAGCCACGTCAA	1067
QY	1749	AGCGGAGAGGGATATCCTAGCAGAGCCGACAAATCAGTGGGTGTCGCCCTGTACTACTC	1808
Db	1068	GGCCGAGAGGGATGTGTTGGCCGAAGCGACTCTCCGTGGGTGCTGCTGTATTACTATTC	1127
QY	1809	TTTCCAGGACAAGGACAACTTGACTTTGTGATGACTACATACTCCTGGGGGGATATGAT	1868
Db	1128	TTTCCAAGATGCCAGTACCTATCTTGATCATGGAATTTTGGCCCGTGGTGACCTGAT	1187
QY	1869	GAGCCTATTAAATTAGATGGGCATCTTTCCTTGAAATCTGGCAGATTTCTACATAGCAGA	1928
Db	1188	GACCATGTTAAATCAGGTGGCAGATATTCACCGAGGACCTCACAGATTTCTACATGGCGGA	1247
QY	1929	ACTTACTCTGCGAGTTGAAAGTGTTCATAAAATGGTTTTATTATCATAGAGATATTAACC	1988
Db	1248	GTGTATCTCGGCATTTAGGCTATACACAGCTGGCTTTTATCCATAGAGATATCAAGCC	1307
QY	1989	TGATAACATTTTGATTGACCGTGATGGCCATATTAAATTGACTGACTTTGGCTTGTGCAC	2048
Db	1308	GGATAACATTCGTGCGACATCAGGGTGCACATCAAACTTTCCGACTTCGCTGTGCGAC	1367
QY	2049	TGCGTTTCAGATGGACACATGACTCCAAGTACTACAGAGTGGGATACACCCACGGCAAGA	2108
Db	1368	AGGGTTCCCAAAACGCGATGACTCCAACTACTACAAGAAGCTGCTTCAGGAGGACGAGCA	1427
QY	2109	TAGCATGGATTTTCAGTAACCAATGGGAGATCCTTCCAATTTGTCGGTGTGG-----	2159
Db	1428	GCAGCAGACGGGGGAACATGGGCAATATCCCGCATCCCGTGGCGGCGCACGGCGG	1487
QY	2160	-----GGACAGACTGAAGCCACTGGA	2180
Db	1488	CGGCAACAGAAACACCATGCTTGTGCGAGCCATCCACCTGACCATGACAAACAGGCAGCA	1547
QY	2181	GGCGAGAGCTGCTCGCCAGCACCCGATGTCTAGCCCATTTCTGTTGGGACTCCCAA	2240
Db	1548	GATGCAAACTCGGCGAAGTCCCGTAGGCTCATGGCTACTTCCACCGTCGGTAGCCGAGA	1607
QY	2241	TTATATTGCACCTGAAGTGCTACTCGCAACAGGATATACACAGCTGTGTACTGGTGGAG	2300
Db	1608	CTACATCGCCCGGAGATCTTCTCTACAGGGCTACGGTCAGGAGTCGCACTGGTGTGTC	1667
QY	2301	TGTTTGGTGTTATTTCTTTGTGAATGTTGGTGGGACACCTCCTTTCTTGGCACAACCCC	2360
Db	1668	CCTCGCGGCATCATGTAGAGTGTCTATCGGGTGGCGCGCTTCTGCTCGGAGACCCC	1727
QY	2361	ATTAGAAACACAAATGAAGTTTATCATCTGGCAAACTTCTCTACACATCCCTCTCAAGC	2420
Db	1728	CCAGGAACCTACAGGAGATCATGAATTTGACGACAGCGTGGTGTTCAGACGACAT	1787
QY	2421	TAAGCTGAGTCTCGAAGCCTCTGACCTCATTTAACTATCAAACTGTCTCGAGGACCCAGAACCG	2480
Db	1788	CCACATCTCATACGAGGACAGGACCTCATCCGCGGCTGCTCTCGCACGCCGACGAAGC	1847
QY	2481	CCTCGCACAAGGGTGTGATGAGATAAAGCTCATCCATTTTTT	2525
Db	1848	GCTCGGTGCGGTACGGCCACAGCAATCAAAACACCCCTTCTT	1892

RESULT 15
US-09-599-256-1.

	Query Match	5.6%	Score 179.8;	DB 4;	Length 2160;
	Best Local Similarity	49.8%	Pred. No. 4.6e-39;		
	Matches 530;	Conservative	0;	Mismatches 487;	Indels 48; Gaps 1;
QY	1509	TGGATTATCTCAAGATGCCCGAGTCAAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAA	1568		
Db	828	TGGCTGTGCTTGAAGAGAGAAGAACACACAGCTTCTTCACTTGGGTAAAAAGGATCGCA	887		

Search completed: January 16, 2003, 21:56:16
Job time : 124.372 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:42 ; Search time 89.1612:Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtgcaacattcaattacg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 39868 seqs, 22934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:**
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	681	21.2	2043	US-09-836-392-2
2	518.6	16.1	638	Sequence 26, Appli
3	376.6	11.7	678	Sequence 26, Appli
4	371.8	11.6	676	Sequence 66, Appli
5	247.6	7.7	1689	Sequence 214, App
6	244.4	7.6	3583	Sequence 1861, Ap
7	212.6	6.6	1452	Sequence 152, App
8	184	5.7	1818	Sequence 2402, Ap
9	175.2	5.5	734	Sequence 89, Appli
10	136.4	4.2	1416	Sequence 196, App
11	132.8	4.1	1244	Sequence 2503, Ap
12	132	4.1	1393	Sequence 36, Appli
13	132	4.1	1735	Sequence 37, Appli
14	130.2	4.1	1398	Sequence 58, Appli
15	117.4	3.7	2365	Sequence 633, App
16	115.8	3.6	2370	Sequence 6, Appli
17	115.8	3.6	2371	Sequence 6, Appli
18	115.8	3.6	2370	Sequence 1, Appli
19	115.8	3.6	2370	Sequence 214, App
				Sequence 3855, Ap

20	113	3.5	1515	10	US-09-804-471A-1	Sequence 1, Appli
21	113	3.5	5877	12	US-10-028-946-3	Sequence 3, Appli
22	113	3.5	6165	12	US-10-028-946-1	Sequence 1, Appli
23	112.8	3.5	1479	10	US-09-771-161A-46	Sequence 46, Appli
24	111.4	3.5	2637	10	US-09-799-875-3	Sequence 3, Appli
25	111.4	3.5	3003	10	US-09-799-875-1	Sequence 1, Appli
26	109.8	3.4	6159	9	US-10-017-216-3	Sequence 3, Appli
27	109.8	3.4	6574	9	US-10-017-216-1	Sequence 1, Appli
28	109.2	3.4	277	10	US-09-294-093B-2043	Sequence 2043, Ap
29	109	3.4	568	9	US-09-796-692-8122	Sequence 8122, Ap
30	109	3.4	5313	10	US-09-801-368-297	Sequence 297, App
31	109	3.4	6409	10	US-09-864-864-293	Sequence 293, App
32	107	3.3	2549	10	US-09-880-107-3691	Sequence 3691, Ap
33	105	3.3	3407	10	US-09-971-845-1	Sequence 1, Appli
34	103.8	3.2	587	9	US-09-764-868-217	Sequence 217, App
35	100	3.1	1194	10	US-09-801-368-399	Sequence 399, App
36	97.2	3.0	2706	12	US-10-071-751-20	Sequence 20, Appli
37	96.4	3.0	2261	9	US-09-954-531-141	Sequence 141, App
38	96.4	3.0	2261	9	US-09-954-531-570	Sequence 570, App
39	96.4	3.0	3061	10	US-09-880-107-2146	Sequence 2146, Ap
40	96.2	3.0	362	9	US-09-796-692-7789	Sequence 7789, Ap
41	96	3.0	6303	9	US-09-974-298-168	Sequence 168, App
42	95.4	3.0	1635	10	US-09-880-107-2340	Sequence 2340, Ap
43	94.2	2.9	512	10	US-09-867-701-6230	Sequence 6230, Ap
44	90	2.8	1197	10	US-09-801-368-403	Sequence 403, App
45	89.8	2.8	2698	10	US-09-749-956-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match 21.2%, Score 681; DB 9; Length 2043;
Best Local Similarity 70.4%; Pred. No. 4.4e-159;
Matches 925; Conservative 0; Mismatches 385; Indels 3; Gaps 1;
QY 1507 GTTGGATTATCTCAAGATCCCGAGGCTAAATGAGAAAGATGCTTTGCCAGAAAGAGTCT 1566
DB 218 GCTGGACTCTGTGAAGCTGAGCAGGAGGAGATGCGGAGAGTCTCTACGAGAAAGTCT 277
QY 1567 AACTATATTCGCTTAAAGGGCTAAATGGACAAGTCTATCTTTGTAAGATAAGACA 1626
DB 278 AATTACACAGGTTAAAGAGGCGCCAGATGAGACAAGTCTATCTTTGTCAAGATCAAAACC 337
QY 1627 TTAGGAATAGAGCGCTTTGGTGAAGTCTGTCTAGCAAGAAAGTCTACTATAAGCTTTG 1686
DB 338 CTGGGGATCGGTGCTTTGGAGAAAGTGTGCCTTGTCTTGAAGTGGACACTCACGCCCTG 397

QY 1687 TATGCAACAAAGACTCTTCCGAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCTCATGTG 1746
 Db 398 TAGCCCATGAAGACCTTAAAGAAAGAGATGCTTGAACCGGAATCAGGTGGCCACGTC 457
 QY 1747 AAGCGGAGAGGATATCTTACGAGAGCGGCAATGAGTGGGTGCTCGCCCTGTACTAC 1806
 Db 458 AAGCGGAGAGGATATCTTACGAGAGCGGCAATGAGTGGGTGCTCGCCCTGTACTAC 517
 QY 1807 TCTTTCCAGGAGGAGCAACTTGTACTTTGTGATGGACTACATCCCTGGGGGGATATG 1866
 Db 518 TCTTTCCAGGAGGAGCAACTTGTACTTTGTGATGGACTACATCCCTGGGGGGATATG 577
 QY 1867 ATGAGCCTAATTAAGAAAGGCACTTCTTCTGAAATCTTGGCAGGATTTCTACATAGCA 1926
 Db 578 ATGAGCCTGCTGATCGGATGAGGCTTCTTCTGAGCAGCTGGCCGCTTCTACATCGCA 637
 QY 1927 GAACCTTACCTGTGCACTTCAAAAGTGTCTATAAATGGGTTTATTTCATAGAGATTTAAA 1986
 Db 638 GAGCTGACTTTGGCCATTGAGAGTGTCCACAAGATGGGCTTCATCCACGAGACATCAAG 697
 QY 1987 CCTGATAACATTTTGAATGACCGTGTATGGCCATATTAATTAAGTACGACTTTGGGCTGTGC 2046
 Db 698 CCTGATAACATTTTGAATGACCGTGTATGGCCATATTAATTAAGTACGACTTTGGGCTGTGC 757
 QY 2047 ACTGGCTTCAGATGACACATGACTTCCAAAGTACTTACCAGAGTGGGGATCCACCGGCAA 2106
 Db 758 ACTGGCTTCAGATGACACATGACTTCCAAAGTACTTACCAGAGTGGGGATCCACCGGCAA 817
 QY 2107 GATAGCTGATTTTCAATTAACAAATGGGAGATCTTCCAAATGCTGCTAGCCATTTCTGTG 2166
 Db 818 GACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 877
 QY 2167 CTGAAGCCTTAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 937
 Db 878 CTGAAGCCTTAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 997
 QY 2227 GTTGGGACTCCCAATTAATTTGCACCTGAAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2286
 Db 937 GTTGGGACTCCCAATTAATTTGCACCTGAAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2346
 QY 2287 TGTGACTGGTGGAGTGTGTTTATTTTGTGAAATGTTTGGTGGGAGCAACCTTCCTTTC 2346
 Db 998 TGTGACTGGTGGAGTGTGTTTATTTTGTGAAATGTTTGGTGGGAGCAACCTTCCTTTC 1057
 QY 2347 TTGGCAGCAACCCATTAAGAAACAAATGAAGTTTATCATCTGGCAAACTTCTCTACAC 2406
 Db 1058 TTGGCAGCAACCCATTAAGAAACAAATGAAGTTTATCATCTGGCAAACTTCTCTACAC 1117
 QY 2407 ATCCCTCTCAAGCTAAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 2466
 Db 1118 ATCCCTCTCAAGCTAAGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTG 1177
 QY 2467 GGACCAAGAGCCCTCGGCAAGAGCGTGTGCTGATGAGATAAAGGCTCATCCATTTTTC 2526
 Db 1178 GGACCAAGAGCCCTCGGCAAGAGCGTGTGCTGATGAGATAAAGGCTCATCCATTTTTC 1237
 QY 2527 AAGACCATGATTTCTTCTAGTGTCTGAGACAGCAGTGTCTTCTATCATCCCTAAATC 2586
 Db 1238 AAGACCATGATTTCTTCTAGTGTCTGAGACAGCAGTGTCTTCTATCATCCCTAAATC 1297
 QY 2587 AGCATCAACAGATATCAATTTTTCAGCCCTGTTGATGATGATGATGATGATGATGATGATGATG 2646
 Db 1298 AGCATCAACAGATATCAATTTTTCAGCCCTGTTGATGATGATGATGATGATGATGATGATGATG 1357
 QY 2647 GGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2706
 Db 1358 GGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1414
 QY 2707 CCGGAG 2766
 Db 1415 CCGGAG 1474

QY 2767 TATAATTATCAAAAGCCTATTGAGTATGAATCAATTCATTCACAGGCGCTCAGA 2819
 Db 1475 TTTGATGCCAAAGCCTTTCAGGAGCAGAAAGCTTCACAGGCTGAGAGCTCAGA 1527
 RESULT 2
 US-09-879-536-26/C
 ; Sequence 26, Application US/09879536
 ; Patent No. US20020144298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/879,536
 ; PRIOR FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/088,801
 ; PRIOR FILING DATE: 1998-06-10
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 26
 ; LENGTH: 638
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-879-536-26

Query Match 16.1%; Score 518.6; DB 10; Length 638;
 Best Local Similarity 89.5%; Pred. No. 4.6e-119;
 Matches 570; Conservative 0; Mismatches 64; Indels 3; Gaps 1;
 QY 1444 GTCTTGAAGTCTCATCAGCAGCCTGTCATCGGAGAGCAGCTAGAAATGAAATGATG 1503
 Db 637 GTACTCAAAATCTCATCAGCAGCCTGTCATCGTAAAAAACAATAGAGAATGAAATGATG 578
 QY 1504 CGGTTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAG 1563
 Db 577 CGGTTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAG 518
 QY 1564 TCTAATATATTCGTTTAAAGGGCTAAATGGACAAGTCTATGTTTGAAGATAAAG 1623
 Db 517 TCTAATATATTCGTTTAAAGGGCTAAATGGACAAGTCTATGTTTGAAGATAAAG 458
 QY 1624 ACATTAAGGATAGGAGCGTTTGGTGAAGTCTCTAGCAAGAAAGTGCATGACTAAAGCT 1683
 Db 457 ACATTAAGGATAGGAGCGTTTGGTGAAGTCTCTAGCAAGAAAGTGCATGACTAAAGCT 398
 QY 1684 TTGTATGCAACAAAGACTCTTCCAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGCTCAT 1743
 Db 397 TTGTATGCAACAAAGACTCTTCCAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGCTCAT 338
 QY 1744 GTGAAGCGGAGGAGGATATCTTAGCAAGCCGACAATGAGTGGGTGGTCCGCTGTAC 1803
 Db 337 GTTAAGGCTGAGAGAGATATCTTGGCTGAAGTGAATGAATGAGTGGTGGTCTATAT 278
 QY 1804 TACTCTTTCCAGGAGGAGGATATCTTAGCAAGCCGACAATGAGTGGGTGGTCCGCTGTAC 1863
 Db 277 TATTCAATCCAAAGATAGGAGCAATTTTACTTTGTAATGGAGTACATCTTCCGCGGTGAT 218
 QY 1864 ATGATGAGCCTATTAAATAGAGTGGGATCTTTCTGAAAAATCTGGCAGGATCTTACATA 1923
 Db 217 ATGATGAGCCTATTAAATAGAGTGGGATCTTTCTGAAAAATCTGGCAGGATCTTACATA 158

QY 1924 GCAGAACTTACCTGTCAGCTTGAAGTGTCTTATATAAATGGTGTATATATAGAGAT 1980
 Db 157 GCAGAACTTACCTGTCAGCTTGAAGTGTCTTATATAAATGGTGTATATATAGAGAT 98
 QY 1981 ATTAACCTGATAACATTTTGTGATGACCTGATGGCCATATTAATAGTACCTTTGGC 2040
 Db 97 ATTAACCTGATAATATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 38
 QY 2041 TTGTGACCTGGCTTCAGATGGACATGACTCCAAGT 2077
 Db 37 CTCGACCTGGCTTCAGATGGACATGACTCCAAGT 1

RESULT 3
 US-09-536-66
 ; Sequence 66, Application US/09879536
 ; Patent No. US20020144298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/879,536
 ; PRIOR FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/088,801
 ; PRIOR FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO: 66
 ; LENGTH: 678
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)....(678)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-879-536-66

Query Match 11.7%; Score 376.6; DB 10; Length 678;
 Best Local Similarity 82.4%; Pred. No. 8.4e-84;
 Matches 500; Conservative 0; Mismatches 97; Indels 10; Gaps 6;
 QY 1447 CTGAAGTCTATCAGCAGCGCTGTCATCGGAAGAAGCAGCTAGAAAATGAAATGATGCGG 1506
 Db 2 CTCAAATCTCATCAGCAGCGCTTACATCGTAAAAACAATAGAGAATGAATGATGCGG 61
 QY 1507 GTTGATATATCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGATCT 1566
 Db 62 GTTGATATATCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGATCT 121
 QY 1567 AACTATATTCGCTTAAAGAGGCTTAAATGGACAAGTCTATCTTGTAAAGATAAAGACA 1626
 Db 122 AATTACATCGCTTAAAGAGGCTTAAATGGACAAGTCTATCTTGTGAAGATAAAGACA 181
 QY 1627 TTAGGAATAGGAGCGTTTGGTGAAGTCTGCTAGCAAGAAAAGTGCATCTAAAGCTTTG 1686
 Db 182 CTAGGAATAGGAGCATTTGGTGAAGTCTGCTAGCAAGAAAAGTGCATCTAAAGCTTTG 241
 QY 1687 TATGCAACAAAGCTCTTCCGAAGAAGACGTTCTGCTCCGAAATCAGGTGCGCTCATGTG 1746
 Db 242 TATGCAACAAAGCTCTTCCGAAGAAGATGTTCTTCTCGAAATCAAGTCTGCTCATGT 301

QY 1747 AAAGCGGAGAGGATATCTTAGCAGAAAGCCGACAAATGAGTGGTGGTCCGCTGTACTAC 1806
 Db 302 AAGGCTGAGAGAGATATCTTGGCTGAAGCTGACATGAATGAGTGGTGGTCTATATAT 361
 QY 1807 TCTTTCCAGGACAAGGACAACCTTTGACTTTGTGATGGAGTACATTCCTGGGGGGATATG 1866
 Db 362 TCATTCGAAGATAAGGGCCATTATCCTT - GTAATGGCTACATTCCTTNGGGGTGATATG 419
 QY 1867 ATGAGCCTATTATAGATGGCATC-TTTCCTGAAATCTGGCAGGATTCATAC--A 1923
 Db 420 AAGAGCCCATTAATANAATGGGCATCTTTCCAGAAAGGCTNGCACCATTCTACCTAG 479
 QY 1924 GCAGAACTTACCTGTG-CAGTTGAAAAGTGTTCAT--AAAATGGGTTTTATTTCATAGAGAT 1980
 Db 480 CCAGAACTTACCTGNGCCNGTTGAAAGTGTCTTAAATGGGTTTAAATCTTAGAGAT 539
 QY 1981 ATTAACCTGATAACATTTTGTGATG--ACCGTGTGAGCCCATATTAATTTGACTGCTTTG 2038
 Db 540 TTTTAACCTGGATAATATTTGANTGCGCGGAAAGGCGCTTATTAATAATGGCTTGCTTTG 599
 QY 2039 GCTTGTG 2045
 Db 600 GCCTTNG 606

RESULT 4
 US-09-764-868-214
 ; Sequence 214, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 214
 ; LENGTH: 676
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (628)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-868-214
 Query Match 11.6%; Score 371.8; DB 9; Length 676;
 Best Local Similarity 73.5%; Pred. No. 1.3e-82;
 Matches 497; Conservative 2; Mismatches 175; Indels 2; Gaps 2;
 QY 1411 TTAAAGTTTTCATGGAGCAGCAGTGTAGAGAAGCTCTGAAAGTCTCATCAGAGCGTCTG 1470
 Db 1 TTAAAGTTTTCATGGAGCAGCAGTGTAGAGAAGTGTATCAAAACCTTACCAGCAGAGGTT 60
 QY 1471 CATCGGAAGAACACAGCTAGAAAATGAATGATGCGGGTGGATTATCTCAAGATGCCAG 1530
 Db 61 AACCGGAGGCTCGAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120
 QY 1531 GATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTTAACATATATTCGCTTAAAGGGCT 1590
 Db 121 GAGCAGATGCGGAAGATGCTCTACCAGAAAGAGTCTAATTACACAGGTTAAAGAGGGCC 180
 QY 1591 AAAATGGACAAGTCTATGTTTGTAAAGATAAAGACATAGGAATAGGAGCGTGTGGTGA 1650
 Db 181 AAGATGGACAAGTCTATGTTTGTCAAGATCAAAACCTTGGGATCGGTGCTTTGGAGAA 240
 QY 1651 GTCTGCTAGCAAGAAAGTGCATTAAGCTTTGTATGCAACAAAGACTCTTCGAAG 1710
 Db 241 GTGTGCTTGTGTTGAAGTGGACACTACGCGCTGTACGCCATGAAGACCCCTAAGGAA 300
 QY 1711 AAAGAGTTTCTCTCCGAAATCAGGTGGCTCATGTGAAGCGGAGGATATCTTAGCA 1770

Fri Jan 17 11:17:00 2003

343 ATGGGGGTTGATGACCTTTGAACCTGTAGCATCATTTGGCGGGTCTTTCGGTGAGGTG 402
 1654 TGTCTAGCAAGAAAGTCGATTAAGCTTTGTATGCAACAAGAGCTCTTCCGAAAGAAA 1713
 403 AGAATTTGTAAGAAAGAAATCTACTGGGAAGCGTATATGCAATGAAAGTTAAAGAAATCC 462
 1714 GACCTTCTGCTCCCAATTCAGTGGCTCATGTGAAAGCGGAGAGGATATCTCTAGCAGAA 1773
 463 GAGATGCTTCCACAGGCGAGTGAACATGTTAAAGCTGAAAGAAATGTCTGTGAGAA 522
 1774 GCCGCAATAGTGGGTGGTCCGCTGTACTACTCTTCTCCAGGCAAGGACAACTTGTAC 1833
 523 GTGATAGCCCTTTCAATGTCAAGCTTTGTACTCTTCTTCCAAAGATGATGACATTTGTAT 582
 1834 TTTGTGATGGACTACATCTCTGGGGGATATGATGAGCCCTTATTAATTAGAATGGGATC 1893
 583 CTATATTTGGAAATACCTCCCTGGAGGTATGATGACACTGCTGTGATCGGAAAGATACC 642
 1894 TTTCTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACTCTGCTGAGTTGAAAGTGT 1953
 643 TTACGGGAAGATGAGACTCGGTTTATCTTGGCAGACAATTTCTGGCTATTTGAGTCTATC 702
 1954 CATAAATGGGTTTATTCATAGAGATATTAACCTGATAACATTTTGTATGACCGTAT 2013
 703 CATAGCAATTAATACGTCCACAGGATATAAGCCTGATAATTTATTGATTACTCGAAAC 762
 2014 GGCCATATTAATTAATGACTGACTTTGGCTGTGCTGCTGCTGAGTGCACATGACTTCC 2073
 763 GGCCATATCAAGCTTTCAATTTGAGTGAAGCAAGTCTCTGGAAAGCAAAATTTTCCA 822
 2074 AGTACTACCAGTGGGATCACCCAGCAAGATAGCATGATTTTCAGTAACCAATGG 2133
 823 GATTTTAAGCGGAGCTTGTGACAGGATACAAGCCCTGCAGCAGACAATGATAGACTC 882
 2134 GGAGATCTCTCAATTTGCTGGTGTGGGACAGACTGAAGCCACTGGAGCGGAGACTGCT 2193
 883 TCCAGGCTCTCTGCACTAG-----GAGAAGCTCAGCAGGAACAGCTTTTACATTGG 936
 2194 CGCCAGCAGCAGGATGCTAGCCCAATTTCTGTTGGGACTCCCAATTTATATTGCACCT 2253
 937 CAACAAACAGAGAGGAGCCCTGGCTTTTCTACAGTAGAAGCTCCCGATTTACATTGCCCT 996
 2254 GAAGTGTCTACTGGCAAGCAGATATACAGCTGTGTGACTGGTGGAGTGTGGTGTATT 2313
 997 GAGTGTCTGCTGAAGAAAGGATATGGAATGGAGTGTGATTTGGTGTCTCTTGAGCAATC 1056
 2314 CTTTGTGAAATGTTGGGACAGCCTCTTCTTGGCAACAAACCCCATTTAGAACACAA 2373
 1057 ATGTTGAGATGCTTTAGGGTTTCCGCCAATTTCTATTCGAGAACTTTGGCAACATGT 1116
 2374 ATGAAGTTATCATCTGGCAAACTTCTTACACATCCCTCCTCAAGCTAGCTGAGTCT 2433
 1117 AGAAGATTGTAACCTGGAACCTGCTTGAAGTTCCCTGATGAAGCTAGCTCTCCATC 1176
 2434 GAAGCTCTGACCTCATTTAATTAACCTGTGCGAGGACCAAGACCCCTCGGCAAGAAC 2493
 1177 GAGTAAAGATCTTATCCGAAGACTGCTCTGCAATGTGCAACAGAGCTTTGGACCAA 1236
 2494 GGTGCTGATGAGTAAAGGCTCATCCATTTTAAAGACCATTCGATTTCTCTAGTGTCTG 2553
 1237 GAGTTCAGAAATCAAGACACACCTTTAGGGAGTCTGAATGGGAAAG--GCTA 1293
 2554 AGACAGCAGTCTGTTTCAATCCCTAAATCAGCATCCCAACAGATACATCCCAATTC 2613
 1294 TATGAGTCAATGCTCCATATATACCAAGTGAAGCATGAGCTGTATACCAAACTTT 1353
 2614 GACCTGTTTGT 2625
 1354 GAAAGATTGAT 1365

RESULT 6
 US-09-974-298-152

301 AAGGATGCTGTAACCGAATAGTGGCCAGCTCAAGCCGAGAGGACATCTGCGCC 360
 1771 GAAGCCGACAATGAGTGGTGGTCCCGTGTACTACTCTTTCCAGGCAAGGACACTTG 1830
 361 GAGCAGACAATGAGTGGTGG-CAAACTCTACTACTCTTCCAGCAAGAGACGCTG 419
 1831 TACTTTGTGAGTACTACATTTCTGGGGGATATGATGAGCCTATTAATTAGAATGGGC 1890
 420 TACTTTGTGAGTACTACATTTCTGGGGGATATGATGAGCCTGCTGATCCGATGGAG 479
 1891 ATCTTTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCGAGTTGAAAGT 1950
 480 GTCCTCTGAGCACTGGCCGKCTACATCGCA-RACTGACTTTGGCCATTTGAGAT 538
 1951 GTTCAATAAATGGGTTTATTCATAGAGATATAAACCCTGATAACATTTTGTGATGACCGT 2010
 539 GTCCACAAGATGGGCTTCATCCAGCAGACATCAAGCCTGATACATTTTGTATGATCTG 598
 2011 GATGCCCATATTAATGACTGACTTTGGCTTTGTCACCTGGCTTTCAGATGGACATGAC 2070
 599 GATGTCATATAACTACAGATTTGCGNCTGCTGACTGGGTTCAAGTGGACTCAAT 658
 2071 TCCAACTACTACCAGA 2086
 659 TTCCAATATTACCAGA 674

RESULT 5
 US-09-938-842A-1861
 ; Sequence 1861, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1861
 ; LENGTH: 1689
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-1861

Query Match 7.7%; Score 247.6; DB 9; Length 1689;
 Best Local Similarity 51.5%; Pred. No. 1.5e-51;
 Matches 624; Conservative 0; Mismatches 579; Indels 9; Gaps 2;
 QY 1414 AAGTTCTTCATGGAGCAGCAGCTAGAGAACGCTCTGAGTCTCATCAGCAGCGCTCTGCAT 1473
 DB 163 AAGCAATATATCGAATCATATATAAATTCAGNAGAAGACTCCAGGAAGAAAGAG 222
 QY 1474 CGGAAGAGCAGCTAGAGAAATGAATGATCGGGTTCGATTTATCTCAAGATGCCAGGAT 1533
 DB 223 CGTCGAGCATCTTGGACAAACCTAGCTGATGCTGATGTTACTGTTGAAGACAAGATG 282
 QY 1534 CAAATGAGAAGAGCTTTGCCAGAAAGCTTAAGTATCTCAAGATGCCAGGAT 1593
 DB 283 GATATATTAAGAACTTTGAGAAAGAAAGAAATGGAGTATATGCTCTACAGACAGAAA 342
 QY 1594 ATGGACAAGCTATCTTTGTTAAAGATAAAGACATTTAGGAATAGAGCGTTGGTGAAGTC 1653

[illegible]

Db 265 ATTAGTCGATGACTTTGAGCTTTGACTATFCAFTGGAAGAGTGCTTTTGGTGAGGTT 324
 QY 1634 TGTCTAGCAAGAAAGTGCATTAAGCTTTGTATGCAACAAAGAGCTCTTCGAAAGAAA 1713
 Db 325 CGCTATCTCGGAGAGAAAGTCTGGAATATTTATGCGATCAAGAACTTAAAGAAATCT 384
 QY 1714 GAGCTTCTGCTCCGAATCAGTGCTCATGTGAAGCGGAGAGGATATCTAGCAGAA 1773
 Db 385 GAAATGGTCATGAGAGGACAGTGTGAGCATGTGAGAGAGAGAGAACTGCTGGCTGAG 444
 QY 1774 GCGGACATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1833
 Db 445 GTTGAAGCCATATATTTGGAAGCTTTTACATTTTCAAGTATCCCGAGTATCTGTAT 504
 QY 1834 TTTGTGATGGACTACATTCCTCGGGGGGATATGATGAGCCTTATTAATAGAAATGGGATC 1893
 Db 505 CTGATTATGGAATATCTCCCGGGTGGTATGATGACCTTGTCTATGAGGAGGACACA 564
 QY 1894 TTTCTGAAATCTGGCAGATTTCTACATGAGCAACTTACCTGTGAGTTGAAAGTGT 1953
 Db 565 TTACGGGAAGATGTTGCCAGATTTTATATGCTCAAGTGTCTGGCCATTTGAATCCATA 624
 QY 1954 CATAAATGGCTTTTATTCATAGAGATATTAACCTGTATACATTTTCAATGACCGTGAT 2013
 Db 625 CACAGATCAACTATTTATAGGATATCAACCTGTATACCTTACCTTTTGGACAAAGAT 684
 QY 2014 GGCCATATTAATGACTGCTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
 Db 685 GGGCACATGAACCTCTCGGACTTTGGCTCTGTAAAGCTCTTGTATGAGAAATTTACCT 744
 QY 2074 AAGTACTACAGAGTGGG-----GATCACCCGCGGAAGATAGCATGGATTTCACT 2124
 Db 745 TCAATTCAGGAGAAATAGGCGCCAGCATGATAACTATGTCAGAACTATGGATTTGAT 804
 QY 2125 AACGAATGGGAGATCTTCAATTTGCTGCTGGGAGAGCTGAGCCACTGGAGCGG 2184
 Db 805 AGATGCTTCTGACACTGATACAGAGAGCTGGGCGAGTCCCGAGCAACAACCTTCAG 864
 QY 2185 AGAGCTGCTGCCAGCAGCAGCTGTCTAGCCCATTTCTGCTGGTGGGACTTCCCAATAT 2244
 Db 865 ---CATTTGGCAGATGAATCGCAGAAACTAGCATTTTCACTGTGGGAACACCGGACTAT 921
 QY 2245 ATTCCACTGAGTGTCTGCGAAGAGATATACAGCTGTGTGACTGGTGGAGTGT 2304
 Db 922 ATTGCTCTGAAGTTTTCCTGGAAGAAAGATATGGCATGGAATGTGATTTGGTGTCTATTA 981
 QY 2305 GGTGTTATTTCTTTGTAATGTTGGTGGGACAACTCTCTTTTGGCACAACCCCATTA 2364
 Db 982 GGTCAATATGACGAATGCTGCTGGTATCCCTTTTATGCTGATGACCTTATA 1041
 QY 2365 GAAACAAATGAAGTTTATCATCTGGCAAACTTCTCTACACATCCCTCTCAAGCTAAG 2424
 Db 1042 TCAACTTGCAGAAAGATGCTCCATTTGGAGAAACCATTTGAATTTCTGTAGGATGCGAAG 1101
 QY 2425 CTGAGTCTGAGCCTCTGACCTCATTTATCAACTGTGTGCGAGGACCCAGAGACCGCTC 2484
 Db 1102 TTTTCATCTGAGGCAAAAGATCTTATCTGCAAGTTGCTGTGCAACGTTGACCATAGGCT 1161
 QY 2485 GGCAAGAACGCTGTGATAGATAAAGGCTTCATCCATTTTAAAGACCATCGATTTCTCT 2544
 Db 1162 GGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
 QY 2545 AGTGATCTGAGACAGAGCTGCTGCTCATACATCCCTTAAATACAGCATCCACAGATACA 2604
 Db 1222 GAAAGGCTCTTATGAAATGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
 QY 2605 TCCAAATTTCCAGCTGTTGATCTCTGATAAT 2635
 Db 1282 CAAATTTTATGAAGTTTGAAGTGAAT 1312

RESULT 8

US-09-771-161A-89
 : Sequence 89, Application US/09771161A
 : Patent No. US20020110811A1
 : GENERAL INFORMATION:
 : APPLICANT: LEVINE, et al.
 : TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
 : FILE REFERENCE: 802620-2005.1
 : CURRENT APPLICATION NUMBER: US/09/771.161A
 : CURRENT FILING DATE: 2001-01-26
 : PRIOR APPLICATION NUMBER: 09/724,676
 : PRIOR FILING DATE: 2000-11-28
 : PRIOR APPLICATION NUMBER: 136776
 : PRIOR FILING DATE: 2000-06-15
 : PRIOR APPLICATION NUMBER: 135619
 : PRIOR FILING DATE: 2000-04-12
 : NUMBER OF SEQ ID NOS: 273
 : SOFTWARE: PatentIn version 3.0
 : SEQ ID NO 89
 : LENGTH: 1818
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : US-09-771-161A-89

Query Match 5.7%; Score 184; DB 10; Length 1818;
 Best Local Similarity 57.1%; Pred. No. 9.6e-36;
 Matches 354; Conservative 0; Mismatches 265; Indels 1; Gaps 1;

QY 1478 AGAAGCAGCTAGAAATGAATGATCGGGTTGGATTATCTCAAGATGCCAGATCAAA 1537
 Db 729 AAAAAGATTAGAAAGGTGATGGAAGAAGAGCCCTAAAGATGAGAGAAACGACTCC 788
 QY 1538 TGAGAAAGATGCTTTGGCAGAAAGAGTCTAACTATATTCGTTCTTAAAGGGCTTAAATGG 1597
 Db 789 GGAGATCAGCACATGCTCGAAGAAACAGAGTTTCTTCGTTTGAAGAGAACAGACTTG 848
 QY 1598 ACAAGTCTATTTTGAAGATAAAGACATTAGAATFAGAGCGTTTGGTGAAGTCTGTG 1657
 Db 849 GATTGGAAGATTTTGAAGTCTCTTAAAGTAATAGGAGAGGAGCATTTTGGTGAAGTACGGC 908
 QY 1658 TAGCAAGAAAGTCGATCTAAAGCTTTGATCAACAAGAGCTCTTCGAAAGAAAGACG 1717
 Db 909 TTGTTCAAGAAAGATACGGGACATGTGATCAATGAAATACCTCCGTAAGACAGATA 968
 QY 1718 TTCTGCTCCGAAATCAGTGGCTTCATGTGAAAGCGGAGAGGATATCTTAGCAGAAAGCGG 1777
 Db 969 TGCTTGAAGAAAGAGCAGGTGGCCACATTCGTCGAGCGGTGACATTTAGTGGAGCAG 1028
 QY 1778 ACAATGAGTGGTGGTCCGCTGTACTCTTTTCCAGGACAGGACAACTTGTACTTTG 1837
 Db 1029 ACAGTTTGGTGGTGTGAAATTTCTATAGTTTTCAGGATAGCTTAAACCTCTACCTAA 1088
 QY 1838 TGATGACTACATTCCTCGGGGGATATGATGAGCCTTATTAATGAAATGGGATCTTTTC 1897
 Db 1089 TCATGAGTTCYTCCTCGGAGGAGCATGATGACCTTTGTATGAAAAAGACA-CTCTG 1147
 QY 1898 CTGAAAATCTGGCAGCATTTCTACATAGCAGAACTTACCTGTGCGAGTTGAAAGTTTCA 1957
 Db 1148 ACAAGAGGAGACTCAAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTTATTCACC 1207
 QY 1958 AAATGGTTTTTATTCATAGATATTAACCTTGAACATTTTGTATTGACCGTGGTGGC 2017
 Db 1208 AACTTGGATTCATCCAGAGACATCAACCCAGACAACTCTTTTGGACCAAGAGGCGC 1267
 QY 2018 ATATTAATTTGACTGCTTTGGCTTGTGACCTTGGCTTTCAGATGGACATGACTTCCAACT 2077
 Db 1268 ATGTGAACTTCTGACTTTGGTCTTTTGCAGAGGACTGAAAAAGACATAGACAGAAAT 1327
 QY 2078 ACTACAGAGTGGGATCAC 2097
 Db 1328 TTTATAGGAATCTGAACCCAC 1347

RESULT 9

RESOURCES
US-09-771-161A-38
; Sequence 38, Application US/09771161A
; Patent No. US2002011081A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROT

838 ACTGATTTTGGTTT 851

Fri Jan 17 11:17:00 2003

```
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724,676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 1244
TYPE: DNA
ORGANISM: Homo sapiens
US-09-771-161A-38

Query Match 4.1%; Score 132.8; DB 10; Length 1244;
Best Local Similarity 58.0%; Pred. No. 4e-23;
Matches 255; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 1609 TTTGTAAGATAAAGACATAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAA 1668
DB 589 TTTGTAATAATTAAGTAATTAAGAGAGGTCTTTTGGTGAGGTGCTGTGTGCAAAATG 648
QY 1669 GTCGATACATAAGCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGA 1728
DB 649 AAGAATACATGAACGAATTTATGCAATGAAATCTCAACAAGTGGGAGATGCTGAAAGA 708
QY 1729 AATCAGGTGGCTCATGTGAAGCGGAGGAGGATATCTTCAGCAAGCGGACATGAGTGG 1788
DB 709 GCAGAGACCGGTCTTCGAGAGAGGCGGATGTCTGGTGAACGGGACTGCCAGTGG 768
QY 1789 GTGGTCCGCTGTACTACTCTTCCAGCAAGGACAACTTTGTACTTTGTGATGGACTAC 1848
DB 769 ATCAGCGCGCTGACATGAGCGGAGGAGGATATCTTCAGCAAGCGGACATGAGTGG 1788
QY 1849 ATTCCTGGGGGGGATGATGAGCGCTTAATTAAGAT---GGGCATCTTCTCTGAAAT 1905
DB 829 TAGTGGGTGGTATTTACTGACCCCTGCTCAGCAAAATTTGAAGACAGCTTCCGGAAGAT 888
QY 1906 CTGGCAGGATTCATACATGACAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGT 1965
DB 889 ATGGCGAGGTCTTACATGTTGTAATGGTGGCCATTGACTTCCATCCATCAGCTTCAT 948
QY 1966 TTTATTATAGAGATATAAACCCTGATACATTTTGAATGACCGGTGATGGCCATATATAA 2025
DB 949 TACGTGCACAGACATTAACCTGACAAATGCTTTTGGAGCGTGAATGGTCATATCCGC 1008
QY 2026 TTGACTGACTTGGCTGTG 2045
DB 1009 CTGGCTGACTTTGGATCATG 1028

RESULT 12
US-09-771-161A-37
Sequence 37 Application US/09771161A
Patent No. US20020110811A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN-KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724,676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 37
```

```
LENGTH: 1393
TYPE: DNA
ORGANISM: Homo sapiens
US-09-771-161A-37

Query Match 4.1%; Score 132.8; DB 10; Length 1393
Best Local Similarity 58.0%; Pred. No. 4.2e-23;
Matches 255; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 1609 TTTGTAAGATAAAGACATAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAA 1668
DB 589 TTTGTAATAATTAAGTAATTAAGAGAGGTCTTTTGGTGAGGTGCTGTGTGCAAAATG 648
QY 1669 GTCGATACATAAGCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGA 1728
DB 649 AAGAATACATGAACGAATTTATGCAATGAAATCTCAACAAGTGGGAGATGCTGAAAGA 708
QY 1729 AATCAGGTGGCTCATGTGAAGCGGAGGAGGATATCTTCAGCAAGCGGACATGAGTGG 1788
DB 709 GCAGAGACCGGTCTTCGAGAGAGGCGGATGTCTGGTGAACGGGACTGCCAGTGG 768
QY 1789 GTGGTCCGCTGTACTACTCTTCCAGCAAGGACAACTTTGTACTTTGTGATGGACTAC 1848
DB 769 ATCAGCGCGCTGACATGAGCGGAGGAGGATATCTTCAGCAAGCGGACATGAGTGG 1905
QY 1849 ATTCCTGGGGGGGATGATGAGCGCTTAATTAAGAT---GGGCATCTTCTCTGAAAT 1905
DB 829 TAGTGGGTGGTATTTACTGACCCCTGCTCAGCAAAATTTGAAGACAGCTTCCGGAAGAT 888
QY 1906 CTGGCAGGATTCATACATGACAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGT 1965
DB 889 ATGGCGAGGTCTTACATGTTGTAATGGTGGCCATTGACTTCCATCCATCAGCTTCAT 948
QY 1966 TTTATTATAGAGATATAAACCCTGATACATTTTGAATGACCGGTGATGGCCATATATAA 2025
DB 949 TACGTGCACAGACATTAACCTGACAAATGCTTTTGGAGCGTGAATGGTCATATCCGC 1008
QY 2026 TTGACTGACTTGGCTGTG 2045
DB 1009 CTGGCTGACTTTGGATCATG 1028

RESULT 13
US-09-764-868-58
Sequence 58 Application US/09764868
Patent No. US20020168711A1
GENERAL INFORMATION:
APPLICANT: ROSEN, et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PTZ32
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 58
LENGTH: 1735
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-868-58

Query Match 4.1%; Score 132; DB 9; Length 1735;
Best Local Similarity 57.5%; Pred. No. 7.6e-23;
Matches 253; Conservative 2; Mismatches 182; Indels 3; Gaps 1;

QY 1609 TTTGTAAGATAAAGACATAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAA 1668
DB 297 TTTGTAATAATTAAGTAATTAAGAGAGGTGCTTTTGGTGAGGTGCTGTGTGCAAAATG 356
QY 1669 GTCGATACATAAGCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGA 1728
DB 357 AAGAATACATGAACGAATTTATGCAATGAAATCTCAACAAGTGGGAGATGCTGAAAGA 416
```


Query Match	4.1%	Score 130.2	DB 9	Length 1398
Best Local Similarity	55.4%	Prod. No. 1.9e-22		
Matches 252	Conservative 0	Mismatches 203	Indels 0	Gaps 0
Qy 1588	GCTAAATGGACAAGTCTTATCTTTGTAAGATAAAGACATAGGAATAGGAGCGTTTGGT	1647		
Db 379	GGTGTAGTGGCATAGATGATTTTGGGTTATGAAGTTTGTGGAAAGGTGCGTTTGGG	438		
Qy 1648	GAAGTCGTCTAGCAAGAAAGTCGATCTAAAGCTTTGTATGTACCAAAAGACTCTTCGA	1707		
Db 439	AAAGTCATACCAGGTGGAGGAAAGGAGACCTCTGAGATATACGCTATGAAAGTCATGAGG	498		
Qy 1708	AAGAAAGACGCTTCCTGCCGAATCAGGTGGCTCATGTGAAAGCGGAGGAGGATTCCTA	1767		
Db 499	AAAGATCACATTATGGAGAAGAACCATGCTGAATACATGAAAGCTGAGCGCGATATTCTT	558		
Qy 1768	GCAGAACCCGACATGAGTGGGTGGTCCGCCCTGTACTACTCTTCCAGGACAAGACAAC	1827		
Db 559	ACAAAAATTGATACCCCATTCATTGTTCAACTTAAATACTCTTTTCAGACAAGTACAGG	618		
Qy 1828	TTGTACTTTGTGTGGACTACATTCCTGGGGGGGATGATGATGAGCGCTATTAAATAGAAATG	1887		

	Query Match	3.7%	Score 117.4	DB 9	Length 2365
	Best Local Similarity	56.5%	Pred. No. 3.8e-19		
	Matches 239	Conservative 0	Mismatches 181	Indels 3	Gaps 1
Qy	1628	TAGGAATAGGAGCGTTTGGTCAAGTCTCTCTAGCAAGAAAGTGCATACATAAACCTTTGT	1687		
Db	371	TCGGAAGGGCGAGTTTGGAAAGGTTCTTCTAGCAAGACACAAAGGCAGAAAGCTGTCT	430		
Qy	1688	ATGCAACAAACACACTCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGGCTCATGTGA	1747		
Db	431	ATGCAGCTAAAGTTTTACAGAAGAAAGCAATCTCTGAAAAGAAAGAGGCAAGCATATTA	490		
Qy	1748	AAGCGGAGAGGGATATCCTAGCAGAAGCCGACAATAG--TGGGTGGTCCGCCCTGTACT	1804		
Db	491	TGTCGGAGCGGAATGTTCTGTTGAAGATGTGAAGCACCCCTTCTCGTGGGCCCTTCACT	550		
Qy	1805	ACTCTTTCCAGGACAAGCAACCTGTACTTTGTATGGAGTACATTCCTCGGGGGGGATA	1864		
Db	551	TCTCTTTCCAGACTGCTGCACAATTTGACTTTGCTCTAGACTACATTAATGGTGGAGAGT	610		
Qy	1865	TGATGAGCCTATTAAATAGATGGCACTCTTCTCGTGAATCTGGCAGATTCTACATAG	1924		
Db	611	TGTTCTACCATCTCCAGAGGAACGCTGCTTCCTGGAAACACGGGCTCGGTTTATGCTG	670		
Qy	1925	CAGAACTTACCTGTGCAGTTGAAAGTGTTTCATAAAATGGGTTTTTATCATAGACATATTA	1984		
Db	671	CTGAATAGCCAGTGCCTTGGGCTACCTGCATTACATGAACATCGTTTATAGACATTAA	730		
Qy	1985	AACCTGATAACATTTTGATTGACCGTGATGCCCATATTAATATGACACTTTTGGCTTGT	2044		
Db	731	AACCAGAGAAATTTTGTCTAGATTACAGGGGACACATTGCTCTTACTGACTTCGGACTCT	790		
Qy	2045	GCA 2047			
Db	791	GCA 793			

Fri Jan 17 11:17:00 2003

us-09-763-334-3.rnpb

Page 10

Search completed: January 16, 2003, 22:03:41
Job time : 133.161 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	671.4	20	9	673	10	BB571112 BB571112
2	653	20	3	665	17	AZ340286 IM0072J17
3	632.8	19	4	667	10	BB174219 BB174219
4	614.4	19	1	648	13	BI155504 602903902
5	567	17	6	650	13	BI693046 603344592
6	548	17	1	600	12	BG081357 H3064C032

ALIGNMENTS

RESULT 1
 BB557112
 LOCUS
 DEFINITION
 BB557112 RIKEN full-length enriched, 2 days pregnant adult female
 ovary Mus musculus cDNA clone E330026M13 3', mRNA sequence.
 BB557112
 ACCESSION
 BB557112.2 GI:16448760
 VERSION
 EST.
 KEYWORDS
 mouse.
 SOURCE
 Mus musculus
 ORGANISM
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 673)
 AUTHORS
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Harai,A.,
 Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
 M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,H.,
 Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
 D., Shibata,K., Shinaawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
 Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
 Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 TITLE
 Unpublished (2001)
 JOURNAL
 On Aug 1, 2000 this sequence version replaced gi:9643478.
 COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center(GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222

Fax: 81-45-503-9216
 E-mail: genome-res@gsc.riken.go.jp,
 URL: <http://genome.gsc.riken.go.jp/>
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 Wachi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES
source[illegible]

QY	1106	CTTCAAGTGTCCCTGTGCATCCCACTCTTCTCTGAAGCTCCAAGCTATCAAGTCCACAC	1165
Db	121	CTTCAAGTGTCCCTGTGCATCCCACTCTTCTCTGAAGCTCCAAGCTATCAAGTCCACAC	180
QY	1166	CGCCTTATCCAAACATCTGCTACCAAAACCCATCTGTCCTCCATATGAGTCAGTAA	1225
Db	181	CGCCTTATCCAAACATCTGCTACCAAAACCCATCTGTCCTCCATATGAGTCAGTAA	240
QY	1226	GTAAGCCCTGCAAGATGAACAGAGCTTACCCAAAGGAGATGATAGTGAGAAGTG	1285
Db	241	GTAAGCCCTGCAAGATGAACAGAGCTTACCCAAAGGAGATGATAGTGAGAAGTG	300
QY	1286	CGGACAGTGGTGACTCTGGGGATAAAGAAACACAGATTACAACCTTCACCTTATCACTG	1345
Db	301	CGGACAGTGGTGACTCTGGGGATAAAGAAACACAGATTACAACCTTCACCTTATCACTG	360
QY	1346	TTCGGAAACAAAGAAAGATGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCCCAC	1405
Db	361	TTCGGAAACAAAGAAAGATGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCCCAC	420
QY	1406	AGGCCTTTAAGTTCTTCATGGAGCAGCAGCTAGAGAAAGCTCTGAAAGTCTCATCAGCAGC	1465
Db	421	AGGCCTTTAAGTTCTTCATGGAGCAGCAGCTAGAGAAAGCTCTGAAAGTCTCATCAGCAGC	480
QY	1466	GTCTGCATCGGAAGACAGCAGCTAGAAAATGAAATGATCGGGTTGGATTATCTCAAGATG	1525
Db	481	GTCTGCATCGGAAGACAGCAGCTAGAAAATGAAATGATCGGGTTGGATTATCTCAAGATG	540
QY	1526	CCGAGGATCAAAATCAGAAAGATGCTTTGCCAGAAAGAGTCAACTATATTCGTCTTAAAA	1585
Db	541	CCGAGGATCAAAATCAGAAAGATGCTTTGCCAGAAAGAGTCAACTATATTCGTCTTAAAA	600
QY	1586	GGGCTAAATGGACAAAGTCTATGTTGTAAGAATAAAGACATAGGAATAGGAGCGTTTG	1645
Db	601	GGGCTAAATGGACAAAGTCTATGTTGTAAGAATAAAGACATAGGAATAGGAGCGTTTG	660
QY	1646	GTGAAGTCTGTCT	1658
Db	661	GTGAAGTCTGTCT	673

RESULT 2

AZ340286/c

LOCUS IM0072117F Mouse 10kb plasmid UUGC1M library Mus musculus genome

DEFINITION IM0072117F. DNA sequence.

665 bp

DNA linear.

GSS 29-SEP

RESULT 2
A7340286/C

RESULT 2
 ACCESSION A2340286
 LOCUS
 DEFINITION
 ACCESSION A2340286
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

High quality sequence stop: 665.

FEATURES
source

1. 665
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0072J17"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (914732114/9b/AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT
ORIGIN

162 a 132 c 204 g 167 t
Query Match 20.3%; Score 653; DB 17; Length 665;
Best Local Similarity 99.8%; Pred No. 9.7e-120;
Matches 664; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 429 GAATCCCCCTAGCCAGGCTCAGAGGCCATTAGTCTCTGTTCCAGTTGGTAGACAA-CCCA 487
DB 665 GAATCCCCCTAGCCAGGCTCAGAGGCCATTAGTCTCTGTTCCAGTTGGTAGACAA-CCCA 606
QY 488 TCATCATCGAGAGTACTAGCAATTTAACTTTACACAGGGGGGACCTGGAGTTCAAGATG 547
DB 605 TCATCATCGAGAGTACTAGCAATTTAACTTTACACAGGGGGGACCTGGAGTTCAAGATG 546
QY 548 GTGGTGGTTCAGTCTCATTTATCTGTCGACCAAAATGTCGCCACTGTTCTGTGACTCGGC 607
DB 545 GTGGTGGTTCAGTCTCATTTATCTGTCGACCAAAATGTCGCCACTGTTCTGTGACTCGGC 486
QY 608 AGCCACCACTCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCTTTACAAA 667
DB 485 AGCCACCACTCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCTTTACAAA 426
QY 668 CAGGGGCTTCTGCTCTCCACCATATTCGCCAATGGAAAGCTTCTCTAGTCGATGATGG 727
DB 425 CAGGGGCTTCTGCTCTCCACCATATTCGCCAATGGAAAGCTTCTCTAGTCGATGATGG 366
QY 728 TGCCCAACAGGAAACAGTACATGAGCTTTAATAATTAACTCTCCCTGGACTGCAAA 787
DB 365 TGCCCAACAGGAAACAGTACATGAGCTTTAATAATTAACTCTCCCTGGACTGCAAA 306
QY 788 CAGCTGCCCCAGTCGCTTCTGCTGCGCAGCTATCCCCAAGCGGTGGGCATGAAA 847
DB 305 CAGCTGCCCCAGTCGCTTCTGCTGCGCAGCTATCCCCAAGCGGTGGGCATGAAA 246
QY 848 TTCCTACATGGCAACCTACATACAGTACAGTCAAAATCTTTTAATACCCATTAGGAA 907
DB 245 TTCCTACATGGCAACCTACATACAGTACAGTCAAAATCTTTTAATACCCATTAGGAA 186
QY 908 GTAGAGCAAGTACATCTGCTTAATCTCAGCTTCTGCGCAGTACAGTCACTGCCATCACAC 967
DB 185 GTAGAGCAAGTACATCTGCTTAATCTCAGCTTCTGCGCAGTACAGTCACTGCCATCACAC 126
QY 968 CCGCTCTCTATTCAACAGCCCGGTGAAAGCATGCGCGTCTCTGAAACAGAGGTGCGAGACTG 1027

Db 125 CCGCTCTCTATTCAACAGCCCGTGAAGAGCATGCGGTCTCTGAACAGAGCTGCGAGACTG 66
QY 1028 CTTTAGCCCCAACCCATCTCTTGTGATGCCACAGCCAGTTTCAGACTGCTTACCTACCC 1087
Db 65 CTTTAGCCCCAACCCATCTCTTGTGATGCCACAGCCAGTTTCAGACTGCTTACCTACCC 6
QY 1088 CTTT 1092
Db 5 CTTT 1
RESULT 3
LOCUS BB174219
DEFINITION BB174219 RIKEN full-length enriched, adult male hypothalamus Mus musculus cDNA clone A230050F21 3' similar to AF104414 Mus musculus large tumor suppressor 1 (Lats1) mRNA, mRNA sequence.
ACCESSION BB174219
VERSION BB174219.2 GI:16269492
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 667).
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
UNPUBLISHED (2001)
On Jun 29, 2000 this sequence version replaced gi:88333302.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp.
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to

FEATURES
 source
 prepare mouse tissues.
 Location/Qualifiers
 1. .667
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A230050F21"
 /clone_lib="RIKEN full-length-enriched, adult male
 hypothalamus"
 /sex="male"
 /tissue_type="hypothalamus"
 /dev_stage="adult"
 /lab_host="DH10B"
 /note="Site 1: Sali; Site 2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN, Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATCCAGAGCTCTTTTNN 3'], cDNA was
 prepared by using trihalose-thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 20.0 and subtraction to Rot = 459.0. Second
 strand cDNA was prepared with the primer adapter of
 sequence [5' GAGAGAGATTCGAGTTAATTAATCCGCCGCCGCC
 3']. cDNA was cleaved with XhoI and BamHI. Vector: a
 modified pBluescript KS(+) after bulk excision from Lambda
 FLC I."
 BASE COUNT 174 a 195 c 142 g 156 t
 ORIGIN
 Query Match 19.4%; Score 622.8; DB 10; Length 667;
 Best Local Similarity 97.0%; Pred No. 9.9e-114;
 Matches 646; Conservative 0; Mismatches 17; Indels 3; Gaps 1;
 QY 641 ATGGACAAGCCCTCTGCTTACAAAGAGGGCTTCTGCTGCCACCATCTTCGCCA 700
 Db 5 AAGGGCAAGCCCTCTGCTTCAACAGAGGGG---TTGTGTGCTCCACCTTCATGGCCA 61
 QY 701 ATGGAAACGTTCTCAGTCGATGATGTCGCCAAGAGAGAGTCATACATGAGCTTT 760
 Db 62 ATGGAAGCGTTCTCAGTCGATGATGTCGCCAAGAGAGAGTCATACATGAGCTTT 121
 QY 761 ATAATATTAAATCTCCTCGATGATGTCGCCAAGAGAGAGTCATACATGAGCTTT 820
 Db 122 ATAATATTAAATCTCCTCGATGATGTCGCCAAGAGAGAGTCATACATGAGCTTT 880
 QY 821 AGTCATCCCCAAGGGTGGGCATGAAATTCCTACATGGCAACCTACATACAGTGGT 241
 Db 182 AGTCATCCACAAGGGTGGGCATGAAATTCCTACATGGCAACCTACATACAGTGGT 241
 QY 881 CAAATTTCTTTAATACCCATTAGGAAGTAGAGCAAGTCACTCTGCTAATTTCTAGCCTT 940
 Db 242 CAAATTTCTTTAATACCCATTAGGAAGTAGAGCAAGTCACTCTGCTAATTTCTAGCCTT 301
 QY 941 CTGCCACTACAGTCAGTCATCAGCCGCTCTTATTCAAGAGCCGCTGAAAGCATGC 1000
 Db 302 CTGCCACTACAGTCAGTCATCAGCCGCTCTTATTCAAGAGCCGCTGAAAGCATGC 361
 QY 1001 GCCTCTGAAACAGAGTCGAGCTGCTTTAGCCCCCAACCCATCTCTTCTGGATGCCAC 1060
 Db 362 GCCTCTGAAACAGAGTCGAGCTGCTTTAGCCCCCAACCCATCTCTTCTGGATGCCAC 421
 QY 1061 AGCCAGTTTCAGCTGTTTACGCTTACCCCTTTTCTGAGGTTACAGCTTCAAGTGGCTG 1120
 Db 422 AGCCAGTTTCAGCTGTTTACGCTTACCCCTTTTCTGAGGTTACAGCTTCAAGTGGCTG 481
 QY 1121 TCATCCCACTGTTGCTGAAGCTCAAGCTATCAAGGTTCCACCGCTTTATCCCAAC 1180
 Db 482 TCATCCCACTGTTGCTGAAGCTCAAGCTATCAAGGTTCCACCGCTTTATCCCAAC 541
 QY 1181 ATCTGCTACACAAACCCATCTGTCCTCCATATGATGATAGTAAGTAAGCCCTGCAAG 1240

Db 542 ATCTGCTACACCAAAACCCATCTGTCCTTCCATATGATGATAGTAAGCCCTGCAAG 601
 QY 1241 ATGAACAGCTAGCTTACCAAGGAAGATGATAGTGAAGAGTCCGACAGTGGTACT 1300
 Db 602 ATGAACAGCTAGCTTACCAAGGAAGATGATAGTGAAGAGTCCGACAGTGGTACT 661
 QY 1301 CTGGGG 1306
 Db 662 CTGGGG 667
 RESULT 4
 BII55504
 LOCUS
 DEFINITION 648 bp mRNA linear EST 05-JUL-2001
 mRNA sequence.
 ACCESSION BII55504
 VERSION BII55504.1 GI:14615505
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 648)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM11093 Row: d Column: 02
 High quality sequence stop: 648.
 Location/Qualifiers
 1. .648
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5033425"
 /clone_lib="NCI_CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."
 BASE COUNT 184 a 132 c 167 g 165 t
 ORIGIN
 Query Match 19.1%; Score 614.4; DB 13; Length 648;
 Best Local Similarity 99.5%; Pred No. 4.7e-112;
 Matches 637; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
 QY 1643 TTGGTCAAGCTCTGTAGCAAGAAAGTCGATCTAAAGC-TTGTATCGACAAAGCTC 1702
 Db 11 TTGGTCAAGCTCTGTAGCAAGAAAGTCGATCTAAAGC-TTGTATCGACAAAGCTC 69
 QY 1703 TTGGAAGAAAGACGCTTCTGCTCCGAAATCAGGTGCTCATGTGAACCGGAGAGGATA 1762
 Db 70 TTGGAAGAAAGACGCTTCTGCTCCGAAATCAGGTGCTCATGTGAACCGGAGAGGATA 128
 QY 1763 TCCTAGCAGAAAGCCCAATGAGTGGGTGGTCCCGCTGTACTCTTTCCAGGACAAGG 1822
 Db 129 TCCTAGCAGAAAGCCCAATGAGTGGGTGGTCCCGCTGTACTCTTTCCAGGACAAGG 188

QY 1823 ACAACTTGTACTTTGTGATGGACTACATCTCTGGGGGGATATGATGAGCCCTTAATAATTA 1882
 Db 189 ACAACTTGTACTTTGTGATGGACTACATCTCTGGGGGGATATGATGAGCCCTTAATAATTA 248
 QY 1883 GAATGGGATCTTCTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCAG 1942
 Db 249 GAATGGGATCTTCTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCAG 308
 QY 1943 TTGAAGTGTTCATAAATGGTTTTATTATCATAGAGATATTAAACCTCATACATTTTCA 2002
 Db 309 TTGAAGTGTTCATAAATGGTTTTATTATCATAGAGATATTAAACCTCATACATTTTGA 368
 QY 2003 TTGACCGTGTGATGGCCATTTAAATTTGACTGCTTTGGCTTTGTGCATCTGGCTTCAGATGA 2062
 Db 369 TTGACCGTGTGATGGCCATTTAAATTTGACTGCTTTGGCTTTGTGCATCTGGCTTCAGATGA 428
 QY 2063 CACATGACTCCAACTACACAGATGGGGATCACCCAGCGGAAGATAGCATGGATTTCA 2122
 Db 429 CACATGACTCCAACTACACAGATGGGGATCACCCAGCGGAAGATAGCATGGATTTCA 488
 QY 2123 GTAAGCAATGGGAGATCTTCCAAATTTGCGGTGTGGGACAGACTGAAGCCACTGGAGC 2182
 Db 489 GTAAGCAATGGGAGATCTTCCAAATTTGCGGTGTGGGACAGACTGAAGCCACTGGAGC 548
 QY 2183 GGAGAGCTGCTCGCCAGCAGCAGCGATGCTAGCCCATTTCTGCTGTGGGACTTCCCAATT 2242
 Db 549 GGAGAGCTGCTCGCCAGCAGCAGCGATGCTAGCCCATTTCTGCTGTGGGACTTCCCAATT 608
 QY 2243 ATATTGACCTGAAGTGTCTACTGCGAAGGATATACACA 2282
 Db 609 ATATTGACCTGAAGTGTCTACTGCGAAGGATATACACA 648

RESULT 5
 BI693046
 LOCUS 603344592F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5372231 5',
 DEFINITION mRNA sequence.
 ACCESSION BI693046
 VERSION BI693046.1 GI:15655675
 KEYWORDS EST.
 SOURCE house;mouse.
 ORGANISM Mus musculus.
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M11947 row: h column: 24
 High quality sequence start: 5
 High quality sequence stop: 643.
 Location/Qualifiers
 1. 650
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5372231"
 /clone_lib="NCI_CGAP_Mam2"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH

BASE COUNT 165 a 198 c 128 g 159 t
 ORIGIN

Query Match 17.6%; Score 567; DB 13; Length 650;
 Best Local Similarity 95.4%; Pred. No. 1.3e-102;
 Matches 629; Conservative 0; Mismatches 20; Indels 10; Gaps 4;

QY 470 CAGTGTGTAGACACCCATCATCTGAGAGTACTAGCAAAATTTAACTTTACACAGGCG 529
 Db 2 CTGTGAGTGTAGACACCCATCATCTGAGAGTACTAGCAAAATTTAACTTTACACAGGCG 55
 QY 530 GACCTGGAGTTCAGAAATGGTGTGCTGCTGATTTATCTGTCACCAAAATGTCGCCA 589
 Db 56 --CGACTGAGTTCAGAAATGGTGTGCTGCTGATTTATCTGTCACCAAAATGTCGCCA 113
 QY 590 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 649
 Db 114 CTGG--TCTGTGACTCGGCAG--CAACACCTCCATATCTCTGACCCCGAGTAATGGACAA 171
 QY 650 GCCCTCTGCTTTACAAACAGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 709
 Db 172 GCCCTCTGCTTTACAAACAGGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 231
 QY 710 TTCTCTCAGTGTGATGTCGCCCAACAGAGCAAGTCAATCAATGAGCTTTATAATATTA 769
 Db 232 TTCTCTCAGTGTGATGTCGCCCAACAGAGCAAGTCAATCAATGAGCTTTATAATATTA 291
 QY 770 ATGTCCTCTGGACTGCAAAACAGCTGCGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
 Db 292 ATGTCCTCTGGACTGCAAAACAGCTGCGCCCGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
 QY 830 CAAGCGGTGGCATGAATTTCTACATGCGCAACCTTAACATACAGTGAAGTCAAAATTCCT 889
 Db 352 CAAGCGGTGGCATGAATTTCTACATGCGCAACCTTAACATACAGTGAAGTCAAAATTCCT 411
 QY 890 TTAATAACCCATTAGGAAGTAGAGCAAGTCACTCTGCTAAATTTCTGAGCTTCTGCGCACTA 949
 Db 412 TTAATAACCCATTAGGAAGTAGAGCAAGTCACTCTGCTAAATTTCTGAGCTTCTGCGCACTA 471
 QY 950 CAGTCACCTGCCATCACCCCGCTCTTATTCACAGCCGCTGAAAGCATGCGCGCTCTGA 1009
 Db 472 CAGTCACCTGCCATCACCCCGCTCTTATTCACAGCCGCTGAAAGCATGCGCGCTCTGA 531
 QY 1010 AACGAGCTGCGAGTGTCTTTAGCCCAACCCATCTCTTCTGGATGCCACAGCAGTTC 1069
 Db 532 AACGAGCTGCGAGTGTCTTTAGCCCAACCCATCTCTTCTGGATGCCACAGCAGTTC 591
 QY 1070 AGACTGTTCAGCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCATCCCA 1128
 Db 592 AGACTGTTCAGCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCATCCCA 650

RESULT 6
 BG081357/c
 LOCUS BG081357
 DEFINITION H3064C02-5 NTA Mouse 15K cDNA Clone Set Mus_musculus cDNA clone
 ACCESSION BG081357
 VERSION BG081357.1 GI:12563925
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 600)
 AUTHORS Kargul, G.J., Dudekula, D.B., Olan, Y., Lim, M.K., Jaradat, S.A., Tanaka
 T.S., Carter, M.G. and Ko, M.S.H.
 TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
 JOURNAL Unpublished (2001)
 COMMENT Other_ESTs: H3064C02-3

Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@gsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3064 row: C column: 02
Seq primer: -21M13 Reverse
High quality sequence stop: 600
POLYA-No.

FEATURES	source
Location/Qualifiers	
1. .600	
/organism="Mus musculus"	
/strain="C57BL/6J"	
/db_xref="niaEST:H3064C02-5"	
/db_xref="taxon:10090"	
/clone="H3064C02"	
/clone_lib="NIA Mouse 15K cDNA Clone Set"	
/sex="Clones arrayed from a variety of cDNA libraries"	
/dev_stage="Clones arrayed from a variety of cDNA libraries"	
/lab_host="DH10B"	
/notes="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos and E12.5 female mesonephros/gonad) and one newborn, ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."	
164	135 g 186 t
BASE COUNT	

Query Match	17.1%	Score 548:	DB 12:	Length 600:	
Best Local Similarity	99.1%	Pred. No. 7.7e-99:			
Matches 551:	Conservative	0:	Mismatches 5:	Indels 0:	Gaps 0:
Qy 2375	TGAAGGTTATCATCTGGCAAACTTCTCTACACATCCCTCCCTCAAGCTAAGCTGAGTCTCTG	2434			
Db 556	TAAAGGTTATCAACTGGCAAACTTCTCTTCACATCCCTCCCTCAAGCTAAGCTGAGTCTCTG	497			
Qy 2435	AAGGCTCTGAGCTCATTTATCAAACTGTCTCAGGACCAGAAAGACCGCTCGCGCAAGACG	2494			
Db 496	AAGGCTCTGAGCTCATTTATCAAACTTGTCTCAGGACCAGAAAGACCGCTCGCGCAAGACG	437			
Qy 2495	GTGCTGATGAGATAAAGGCTCATCCATTTTTTAAGACCATCGATTTCTCTAGTGCATCTGA	2554			
Db 436	GTGCTGATGAGATAAAGGCTCATCCATTTTTTAAGACCATCGATTTCTCTAGTGCATCTGA	377			
Qy 2555	GACAGCAGTCTGCTTCATACATCCCTAAAATCAGCATCCAAACAGATACATCCCAATTTTCG	2614			
Db 376	GACAGCAGTCTGCTTCATACATCCCTAAAATCAGCATCCAAACAGATACATCCCAATTTTCG	317			
Qy 2615	ACCCTGTTTGATCCCTGATAAAATTTGGAGCGGATGGCAGCGAGGAGGAAATATCAGTGGACA	2674			
Db 316	ACCCTGTTTGATCCCTGATAAAATTTGGAGCGGATGGCAGCGAGGAGGAAATATCAGTGGACA	257			
Qy 2675	CTCTGAGCGGGATGGTATAAATAATGGGAAGCACCCCGAGCAGCGCTTCTCTATGAGTTCACTT	2734			
Db 256	CTCTGAACGGATGGTATAAATAATGGGAAGCACCCCGAGCAGCGCTTCTCTATGAGTTCACTT	197			

QY	2735	TTGGAGGGTTTTTGATGACAATGGCTACCCTATATATTCACAAAGCCCTATTGAGTATG	2794
Db	196	TTGGAGGGTTTTTGATGACAATGGCTACCCTATATATTCACAAAGCCCTATTGAGTATG	137
QY	2795	AATACATTCAATTCACAGGGCTCAGAACACAGCTGATGAAGATGATCAACACACAAGCT	2854
Db	136	AATACATTCAATTCACAGGGCTCAGAACACAGCTGATGAAGATGATCAACACACAAGCT	77
QY	2855	CCGATGGAACCAACCGAGATCTAGTGTATGTTTAATAAAGTAGGAGATCATTTGTAAGAAT	2914
Db	76	CCGATGGAACCAACCGAGATCTAGTGTATGTTTAATAAAGTAGGAGATCATTTGTAAGAAT	17
QY	2915	TTGCAAGAGGCGCTGAA	2930
Db	16	TTGCAAGAGGCGCTGAA	1
RESULT 7			
LOCUS	BG081236/c		
DEFINITION	BG081236	621 bp mRNA linear	EST 26-JAN-2001
	H3062G212-5	NIA Mouse 15K cDNA Clone Set	Mus musculus cDNA clone
	H3062G212 5'	mRNA sequence.	
ACCESSION	BG081236		
VERSION	BG081236.1	GI:12563804	
KEYWORDS	EST.		
SOURCE	house mouse		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 621)		
AUTHORS	Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Jaradat, S.A., Tanaka		
	T.S., Carter, M.G. and Ko, M.S.H.		
TITLE	Verification and Initial annotation of NIA mouse 15K cDNA clone set		
JOURNAL	Unpublished (2001)		
COMMENT	Other ESTs: H3062G12-3		
	Contact: George J. Kargul		
	Laboratory of Genetics		
	National Institute on Aging/National Institutes of Health		
	333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA		
	Email: cdnaelg@nigms.nih.gov		
	This clone set has been freely distributed to the community. Please		
	visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.		
	Plate: H3062 Row: G Column: 12		
	Seq primer: -21M13 Reverse		
	High quality sequence stop: 621		
	POLYA=No.		

```

FEATURES
source
Location/Qualifiers
1. 621
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3062G12-5"
/db_xref="taxon:10090"
/clone="H3062G12"
/clone_lib="NIA Mouse 15K cDNA Clone Set"
/sex="Clones arrayed from a variety of cDNA libraries"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/notes="vector: pSPORT1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos , and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dT) -Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of

```

unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 169 a 120 c 139 g 193 t
ORIGIN

Query Match 17.1%; Score 548; DB 12; Length 621;
Best Local Similarity 99.1%; Pred. No. 7.6e-99;
Matches 551; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2375 TGAAGGTTATCATCTGGCAAACTTCTTACACATCCCTCCTCAAGCTAGCTGAGTCGCG 2434
DB 556 TAAAGGTTATCACTGGCAAACTTCTTACACATCCCTCCTCAAGCTAGCTGAGTCGCG 497
QY 2435 AAGCCTCTGACCTCATATCAACCTGTCGAGGAGCCAGAACCCCTCGGCAAGAGC 2494
DB 496 AAGCCTCTGACCTCATATCAACCTTTCGAGGAGCCAGAACCCCTCGGCAAGAGC 437
QY 2495 GTGCTGATGAGATAAGGCTCATCCATTTTAAAGACCATCGATTCTCTAGTGATCTGA 2554
DB 436 GTGCTGATGAGATAAGGCTCATCCATTTTAAAGACCATCGATTCTCTAGTGATCTGA 377
QY 2555 GACAGAGTCTGCTTATACATCCCTTAAATCACGCATCCCAACAGATACATCCAAATTCG 2614
DB 376 GACAGAGTCTGCTTATACATCCCTTAAATCACGCATCCCAACAGATACATCCAAATTCG 317
QY 2615 ACCCTGTTGATCCTGATTAATTTGTGGAGGATGGGAGGAGGAGGAGGAGGAGGAGGAG 2674
DB 316 ACCCTGTTGATCCTGATTAATTTGTGGAGGATGGGAGGAGGAGGAGGAGGAGGAGGAG 257
QY 2675 CTCTGAGGGATGGTATAAAATGGAGACCCCGAGCAGCCTTCTATGAGTTCACCT 2734
DB 256 CTCTGAGGGATGGTATAAAATGGAGACCCCGAGCAGCCTTCTATGAGTTCACCT 197
QY 2735 TTCGAGGTTTTGATGACAAATGGCTACCCATATAATTAATTAATTAATTAATTAATTAAT 2794
DB 196 TTCGAGGTTTTGATGACAAATGGCTACCCATATAATTAATTAATTAATTAATTAATTAAT 137
QY 2795 AATACATTCATTCACAGGCTCAGAACACACAGCTGATGAAGATGATCAACACAGCT 2854
DB 136 AATACATTCATTCACAGGCTCAGAACACACAGCTGATGAAGATGATCAACACAGCT 77
QY 2855 CCGATGGAACACCGAGATCTAGTGTATGTTTAAATAAATAGGAGATCATTTGTAAGAT 2914
DB 76 CCGATGGAACACCGAGATCTAGTGTATGTTTAAATAAATAGGAGATCATTTGTAAGAT 17
QY 2915 TTGCAAGAGGCTGAA 2930
DB 16 TTGCAAGAGGCTGAA 1

RESULT 8
AV708169 LOCUS 694 bp mRNA linear EST 09-OCT-2000
DEFINITION AV708169 ADC Homo sapiens cDNA clone ADCALE11 5', mRNA sequence.
ACCESSION AV708169
VERSION AV708169.1 GI:10725434
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 694)
REFERENCE Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA ADC clones
Unpublished (2000)
JOURNAL
COMMENT Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai

201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgehc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
Location/Qualifiers
1..694
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADCALE11"
/clone_lib="ADC"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 229 a 177 c 124 g 162 t 2 others
ORIGIN

Query Match 15.8%; Score 508.4; DB 10; Length 694;
Best Local Similarity 85.0%; Pred. No. 5.7e-91;
Matches 591; Conservative 0; Mismatches 102; Indels 2; Gaps 2;

QY 743 GTCATAACATGAGGCTTTATATATTAATGTCCTCGACTGCAACAGCCTGGCCCCAGT 802
DB 1 GTCATAACATGAGGCTTTATATATTAATGTCCTCGACTGCAACAGCCTGGCCCCAGT 60
QY 803 CGTCTTCTGCTCCTGGGAGTATCCCAAGGGTGGGATGAAATTCCTACATGGCAAC 862
DB 61 CATCTTCTGCTCCAGCCCATCATCCCGAGCAGTGGGATGAAATTCCTACATGGCAAC 120
QY 863 CTAACATACCATGAGTCAAAATTCCTTTAATACCCATTAGGAAGTAGCAAGTCACT 922
DB 121 CTAACATACCATGAGTCAAAATTCCTTTAATACCCATTAGGAAGTAGCAAGTCACT 180
QY 923 CTGCTAAATTCAGGCTTCTGCCATACAGTCACTGCCATCACCAGCTCCCTATTTCAAC 982
DB 181 CTGCTAAATTCAGGCTTCTGCCATACAGTCACTGCCATCACCAGCTCCCTATTTCAAC 240
QY 983 AGCCGCTGAAAGCATGCGCGTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCCAACCC 1042
DB 241 AGCCGCTGAAAGCATGCGCGTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCCAACCC 300
QY 1043 ATCTCTTCTGGATGCCACAGCCAGTTCAGACTGTTTCAGCCTACCCCTTTTCTGAGGGTA 1102
DB 301 ACCCTTCTGGATACCACAGCCAAATTCAACTGTTCAACCCAGTCTCTTTCTGAGGGAA 360
QY 1103 CAGCTTCAAGTGTGCTGTCATCCCACTGTTGCTGGAAGCTCCAGCTATCAAGGTCCAC 1162
DB 361 CCGCTTCAAAATGCTGACTGTGATGCCACCTGTTGCTGGAAGCTCCAACTATCAAGGACCAC 420
QY 1163 CACCGCTTATCCAAACATCTGTACACCAAAACCCATCTGCTCCCTCATATGAGTCAG 1222
DB 421 CACCGCTTATCCAAACATCTGTGTCACCAAAACCCATCTGCTCCCTCATATGAGTCAG 480
QY 1223 TAACTAAGCCCTGCAAGATGAACAGCTTAGCTTACCCAGGAAGATGATGATGAGAAGA 1282
DB 481 TCAGTAAGCTTAGCAAGAGGATCAGCCAGCTTGGCCCAAGGAAGATGAGATGAAAGA 540
QY 1283 GTGCGGACAGTGTGCTGCTGGGATTAAGAAAGAAAGAAAGATTAACAAGTTCACCTATCA 1342
DB 541 GTTATGAAATCTTGTAGTGGGATTAAGAAAGAAAGAAAGATTAACAAGTTCACCTATTA 599
QY 1343 CTGTTGCGGAAACCAAGAGATGAAGACGAGAGAGTCTGGATTAGAGTTACTCTCC 1402
DB 600 CTGTTGCGGAAACCAAGAGATGAAGACGAGAGAGTCTGGATTAGAGTTACTCTCC 659
QY 1403 CACAGGCTTTAAGTTCCTCATGGAGCAGCAGTA 1437
DB 660 CTCAGAG-CATTAAATCTTTATGGAGCCACATGTA 693

RESULT 9

BQ201853/c
 LOCUS 602 bp mRNA linear EST 02-MAY-2002
 DEFINITION UI-R-D01-cml-b-11-0-UI.s1 UI-R-D01 Rattus norvegicus cDNA clone
 ACCESSION BQ201853
 VERSION BQ201853.1 GI:20418318
 KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 602)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 The sequence contained an oligo-dT track that was present in the
 oligonucleotide that was used to prime the synthesis of first
 strand cDNA and therefore this may represent a bonafide poly A
 tail. The sequence tag present in the cDNA between the NOL1 site
 and the oligo-dT track served to verify it as a clone from the
 normalized rat cell line R3327-5p library cDNA library Preparation:
 M.B. Soares Lab Clone distribution: clones will be available
 through Research Genetics (www.resgen.com)
 Seq primer: M13 Forward
 POLYA=Yes.

FEATURES

Location/Qualifiers
 1..602
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-D01-cml-b-11-0-UI"
 /clone_lib="UI-R-D01"
 /dev_stage="adult"
 /lab_host="Dhl0B (Life Technologies)"
 /notes="vector: pT7R3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The UI-R-D01
 library is a normalized Rat Cell Line R3327-5p library
 (nR5P) constructed in pT377 PAC vector according to the
 procedure described by Bonaldo, Lennon & Soares
 (Normalization and Subtraction: Two Approaches to
 Facilitate Gene Discovery. Genome Research 6: 791-806,
 1996.) The oligonucleotide used to prime first strand
 synthesis contained the sequence tag CACGTGAGAT between
 the Not I cloning site and dT18 stretch. The rat cell line
 R3327-5p was provided by Mary Hendrix of the University of
 Iowa.
 TAG_LIB=UI-R-D01
 TAG_TISSUE=rat cell line R3327-5p
 TAG_SEQ=CACGTGAGAT"

BASE COUNT 176 a 120 c 108 g 198 t
 ORIGIN

Query Match 15.5%; Score 497.2; DB 14; Length 602;
 Best Local Similarity 90.9%; Pred. No. 9.8e-89;
 Matches 540; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

QY 2513 CTCATCCATTTTAAAGACCATGATTTCTAGTGATCTGAGACAGAGCTGCTTCAT 2572
 |||||||
 DB 602 CTCATCCATTTTAAAGACCATGATTTCTAGTGATCTGAGACAGAGCTGCTTCAT 543
 |||||||
 QY 2573 ACATCCCTAAATACCGCATCCACAGATACATAATTTGACCCCTGTTGATCCTGATA 2632
 |||||||
 DB 542 ACATCCCAAAATACCGCACCCACACAGACAGCTCCCAATTTTGACCCCTGTTGACCCCTGATA 483

QY 2633 AATTGTGGAGCGATGGCAGCGAGGAGGAGAAATATATCAGTCACACTCTGACGGAGTGATATA 2692
 |||||||
 DB 482 AGTTGTGGCGCGATGTTAAACGAGGAGAGAAATATCAACGACACTCTGAAACGGATGATATA 423
 |||||||
 QY 2693 AAAATGGGAAGACACCCCGAGCAGCGCTTTCTATGAGTTCCACTTTCGGAGGTTTGTGATG 2752
 |||||
 DB 422 AAAACGGAAGCACCCTGAGCATGCTTTGTATGAGTTTACCTTTTCGGAGGTTCTTCTGATG 363
 |||||
 QY 2753 ACAATGGCTACCCATATATATATCCAAAGCCTATTGAGTATGAATACATTCATTCACAGG 2812
 |||||
 DB 362 ACACTGGCTACCCATATATATATCCAAAGCCTATCAATATGATATAGTTAGTTCCAGG 303
 |||||
 QY 2813 GCTCAGACAACACAGTCTGATGAAGATGATCAACACACAGCTCCGATGGAACACCGAG 2872
 |||||
 DB 302 GCTCAGACAGCAGCTCCGATGAAGACGATCAACACACAGCTCAGATGTAACACACCGTG 243
 |||||
 QY 2873 AYTCTAGTATGTTTAACTAACTAGGAGATCAATGTAAAGAAATTTGCA-AGAGCCCTGAAG 2931
 |||||
 DB 242 ATCTAGTATGTTTAACTAACTAGGAGATCAATGTAAAGAAATTTGCAAGAGCCCTGAAA 183
 |||||
 QY 2932 TGCAGGGGTTTTTGAAGTTTTTGAAGAAATTTATGCAATGTGACAGAGTTTGTGCTGCTG 2991
 |||||
 DB 182 TGCAGGGGTTTTTGAAGTTTTTGAAGAAATTTATGCAATGTGACAGAGTTTGTGCTGCTG 123
 |||||
 QY 2992 TGTACAAATATTTTATTTTCTCTAAGTTATGGGAATTTGTTTAAATGTTTAAATTTATTC 3051
 |||||
 DB 122 TGTACAAATATTTTATTTTCTCTAAGTTATGGGAATTTGTTTAAATGTTTAAATTTATTC 63
 |||||
 QY 3052 CCCTTTTAACTAGTAAATTTAGAAAAATTTGTTATAGGAAAGTAAATTTATGAA 3105
 |||||
 DB 62 CCCTTTTAACTAGTAAATTTAGAAAAATTTGTTATAGGAAAGTAAATTTATGAA 9

RESULT 10
 BE093083 578 bp mRNA linear EST 12-JUN-2000
 LOCUS RC5-BT0744-260400-031-G07 BT0744 Homo sapiens cDNA, mRNA sequence.
 DEFINITION BE093083
 ACCESSION BE093083
 VERSION BE093083.1 GI:8483535
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 578)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjowski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
 M.J., Soares,F., Brentani,R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 MEDLINE 20202663
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-RC5-BT0744-260
 400-031-G07&t3=2000-04-26&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 12
 High quality sequence stop: 467.
 Location/Qualifiers
 1..578

FEATURES

source 1..578

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="B70744"
/dev_stage="Adult"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 163 a 120 c 129 g 165 t 1 others
ORIGIN

Query Match 15.1%; Score 486.4; DB 10; Length 578;
Best Local Similarity 91.6%; Pred. No. 1.4e-86;
Matches 514; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1871 GCCTATTAAATGAATGGGCACTTTCTGAAAATCTGCACGATTCTACATAGAGAAC 1930
DB 18 GCCTATTAAATGAATGGGCACTTTCTCAGAAAGTCTGCACGATTCTACATAGAGAAC 77

QY 1931 TTACCTGTGCAGTTCGAAGTGTTCATAAATGGGTTTATTTCATAGAGATATTAACCTG 1990
DB 78 TTACCTGTGCAGTTCGAAGTGTTCATAAATGGGTTTATTTCATAGAGATATTAACCTG 137

QY 1991 ATAACATTTTGAATGACCGTGTATGCCATATTAATAATGACTGCTTGGCTTGTGCACGTG 2050
DB 138 ATAATATTTTGAATGCTGTGATGCTATTAATAATGACTGCTTGGCTTGTGCACGTG 197

QY 2051 GCTTCAGATGGACATGACTCAAGTACTACAGAGTGGGATCACCCACGCGCAAGATA 2110
DB 198 GCTTCAGATGGACATGACTCAAGTACTACAGAGTGGGATCACCCACGCGCAAGATA 257

QY 2111 GCATGGATTTAGTAACGAATGGGAGATCCCTTCCAAATGTCGGTGTGGGACACACTGA 2170
DB 258 GCATGGATTTAGTAATGAATGGGAGATCCCTCAAGTGTGGTGTGGGACACACTGA 317

QY 2171 AGCCACTGGAGGGAGAGTGTCTGCCAGACGACGATGTCTAGCCATTTCTGTGGTGG 2230
DB 318 AGCCATTAGAGGGAGAGTGTCTGCCAGACGACGATGTCTAGCCATTTCTGTGGTGG 377

QY 2231 GGACTCCCAATTTATATTCACCTGAAGTGTCTACTCGCAACAGGATATACACAGCTGTGTG 2290
DB 378 GGACTCCCAATTTATATTCACCTGAAGTGTCTACTCGCAACAGGATATACACAGTGTGTG 437

QY 2291 ACTGTGTGAGTTCGGTGTATTCTTTGTGAATGTTGGTGGGACAACTCCTTTCTTGG 2350
DB 438 ATTGTGTGAGTTCGGTGTATTCTTTGTGAATGTTGGTGGGACAACTCCTTTCTTGG 497

QY 2351 CACAAACCCATTAGAAACAAATGAAGTGTATCATCTGCAAACTTCTCTACACATCC 2410
DB 498 CACAAACCCATTAGAAACAAATGAAGTGTATCATCTGCAAACTTCTCTACACATCTCTTACATTC 557

QY 2411 CTCCTCAAGCTAAGCTGAGTC 2431
DB 558 CACCACAAAGCTAAACTCAGTC 578

RESULT 11
AI594372
LOCUS
DEFINITION
IMAGE:1049993 5' similar to TR:Q24096 Q24096 LATS. [1] ; mRNA
sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AI594372
AI594372.1 GI:4603420
EST
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 496)

AUTHORS

Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, F., Jackson, I., Cardenas, M., McCann, R., Waterston, R., and Willson, R.

The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

Contact: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:581569

This read is a RESEQUENCE of a previously sequenced mouse clone correct orientation

Seq primer: -40RP from Gibco

High quality sequence stop: 472.

FEATURES

source

Location/Qualifiers

1..496

/organism="Mus musculus"

/strain="C3H"

/db_xref="taxon:10090"

/clone="IMAGE:1049993"

/clone_lib="Barstead mouse myotubes MPLRB5"

/cell_line="C2C12"

/lab_host="DH10B"

/note="Vector: pT73B-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a NotI oligo(dT) primer [5' TGTACGAATCTGAAGTGGAGCGCCCTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob Barstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."

BASE COUNT 138 a 96 c 129 g 132 t 1 others

ORIGIN

Query Match 15.0%; Score 481.4; DB 9; Length 496;
Best Local Similarity 99.6%; Pred. No. 1.4e-85;
Matches 482; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1701 TCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGGGA 1760
DB 13 TCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGGGA 72

QY 1761 TATCCTAGCAGAACCGGACAAATGAGTGGTGTCCGCTGTACTACTCTTTCCAGACAA 1820
DB 73 TATCCTAGCAGAACCGGACAAATGAGTGGTGTCCGCTGTACTACTCTTTCCAGACAA 132

QY 1821 GGACAACCTGTACTTTGTGATGAGTACATTCCTCGGGGGATATCATGAGCTTAAAT 1880
DB 133 GGACAACCTGTACTTTGTGATGAGTACATTCCTCGGGGGATATGATGAGCTTAAAT 192

QY 1881 TAGAATGGGCACTTTCTCTGAAAATCTGCACGATTTACATAGCAGAACTTACCTGTGC 1940
DB 193 TAGAATGGGCACTTTCTCTGAAAATCTGCACGATTTACATAGCAGAACTTACCTGTGC 252

QY 1941 AGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGAGATATTAACCTGTAAACATTTT 2000
DB 253 AGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGAGATATTAACCTGTAAACATTTT 312

QY 2001 GATTGACCGTGTATGCCATATTAATTAATGACTGACATTTGGCTTGTGCATGGCTTCAGATG 2060
DB 313 GATTGACCGTGTATGCCATATTAATTAATGACTGACATTTGGCTTGTGCATGGCTTCAGATG 372

QY 2061 GACACATGACTCCAAGTACTACACAGTGGGATCACCCAGCGAAGATAGATGATGATTT 2120

was primed with a Not I - oligo(dT) primer [5']
 TGTACCAATCTGAAGTGGAGCGCGCGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
 State Univ., from 2); double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT773 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bonaldo.

BASE COUNT 117 a 142 c 101 g 113 t
 ORIGIN
 Query Match 14.6%; Score 468.2; DB 9; Length 473;
 Best Local Similarity 99.4%; Pred. No. 6e-83;
 Matches 470; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 317 AGACAACCGCTACTCTCGGACATGAGTACGTAATCTCCGAAATCTCCCTGTTCAC 376
 Db 1 AGTCAACCGCTACTCTCGGACATGAGTACGTAATCTCCGAAATCTCCCTGTTCAC 60
 Qy 377 CTGGGCGTGCAGGAGGCTACCTCCACCACTCTTACCACCTCTCCCATGATCCCC 436
 Db 61 CTGGGCGTGCAGGAGGCTACCTCCACCACTCTTACCACCTCTCCCATGATCCCC 120
 Qy 437 CTAGCAGGCTCAGAGGCGCCATTAGTCTGTTCAGTGGTAGACAACCCATCATGTC 496
 Db 121 CTAGCAGGCTCAGAGGCGCCATTAGTCTGTTCAGTGGTAGACAACCCATCATGTC 180
 Qy 497 AGAGTACTAGCAATTTAACTTTACACGAGGCGACCTGGAGTTTCAGAAATGGTGTGTC 556
 Db 181 AGAGTACTAGCAATTTAACTTTACACGAGGCGACCTGGAGTTTCAGAAATGGTGTGTC 240
 Qy 557 AGTCTGATTTATCTGTCGACCAAAATGTCACCACTGTTCTGTGACTCGGACGACCAAC 616
 Db 241 AGTCTGATTTATCTGTCGACCAAAATGTCACCACTGTTCTGTGACTCGGACGACCAAC 300
 Qy 617 CTCCATATCTCTGACCCGCTAATGGACAAGCCCTCTGCTTTTACAAACAGGGGCTT 676
 Db 301 CTCCATATCTCTGACCCGCTAATGGACAAGCCCTCTGCTTTTACAAACAGGGGCTT 360
 Qy 677 CTGCTGCTCCACCATCATTCGCAATGGAAGAGTTCCTCAGTGCATGATGTTGCCAACA 736
 Db 361 CTGCTGCTCCACCATCATTCGCAATGGAAGAGTTCCTCAGTGCATGATGTTGCCAACA 420
 Qy 737 GGAACAGTCATAACATGAGCTTTATATATTATGTCCTCCCTGGACTGCAACA 789
 Db 421 GGAACAGTCATAACATGAGCTTTATATATTATGTCCTCCCTGGACTGCAACA 473

RESULT 14
 AI596344/c 458 bp mRNA linear EST 21-APR-1999
 LOCUS
 DEFINITION
 me57h09.x1 Soares mouse embryo NME13.5 14.5 Mus musculus cDNA
 clone IMAGE:391649 3' similar to TR:Q24096 Q24096 LATS. [1] ;, mRNA
 sequence.
 ACCESSION
 VERSION AI596344.1 GI:4605392
 KEYWORDS
 SOURCE EST.
 ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 458)
 REFERENCES
 AUTHORS
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person,
 B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter,
 E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterston, R., and Wilson, R.
 TITLE
 The Washu-NCI Mouse EST Project 1999
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Marra M/Washu-NCI Mouse EST Project 1999
 Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.lnl.gov) for further information.
 This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 High quality sequence stop: 449.

FEATURES

Location/Qualifiers
 1..458
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:391649"
 /clone_lib="Soares mouse embryo NME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 /note="Vector: pT773D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5']
 TGTACCAATCTGAAGTGGAGCGCGCGAAATTTTTTTTTTTTTTTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
 State Univ., from 2); double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT773 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M. Fatima Bonaldo."
 BASE COUNT 98 a 98 c 119 g 143 t
 ORIGIN

Query Match 14.2%; Score 455; DB 9; Length 458;
 Best Local Similarity 100.0%; Pred. No. 2.5e-80;
 Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1039 ACCCATCTCTTCTGGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTTCTCAG 1098
 Db 458 ACCCATCTCTTCTGGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTTCTCAG 399
 Qy 1099 GGTACAGCTTCAAGTGTGCCTGTATCCACCTGTGCTGAAGCTCCCAAGCTATCAAGT 1158
 Db 398 GGTACAGCTTCAAGTGTGCCTGTATCCACCTGTGCTGAAGCTCCCAAGCTATCAAGT 339
 Qy 1159 CCACCACCGCTTATCCAAAACATCTGTACACCAAAACCCATCTGTCCCTCATATGAG 1218
 Db 338 CCACCACCGCTTATCCAAAACATCTGTACACCAAAACCCATCTGTCCCTCATATGAG 279
 Qy 1219 TCAGTAAGTAAAGCCCTGCAAGATGAACAGCCCTAGCTTACCCAGAGAGATGATGAG 1278
 Db 278 TCAGTAAGTAAAGCCCTGCAAGATGAACAGCCCTAGCTTACCCAGAGAGATGATGAG 219
 Qy 1279 AAGAGTGGCGACAGTGTGACTCTGGGATAAGAAAAGAACAGATTACAACCTTCACT 1338
 Db 218 AAGAGTGGCGACAGTGTGACTCTGGGATAAGAAAAGAACAGATTACAACCTTCACT 159
 Qy 1339 ATCACTGTTTCGAAAAACAAGAAAGATGAAGAACGAGAGTCTCGGATTTCAGAGTTAC 1398
 Db 158 ATCACTGTTTCGAAAAACAAGAAAGATGAAGAACGAGAGTCTCGGATTTCAGAGTTAC 99
 Qy 1399 TCCCCACAGGCTTTAAGTCTTCTATGAGCAGCAGCAGCTAGAGAACGCTCCTGAAGTCTCAT 1458
 Db 98 TCCCCACAGGCTTTAAGTCTTCTATGAGCAGCAGCAGCTAGAGAACGCTCCTGAAGTCTCAT 39
 Qy 1459 CAGCAGCCTCTGCATCGGAAGAACAGCAGCTAGAAAA 1493
 Db 38 CAGCAGCCTCTGCATCGGAAGAACAGCAGCTAGAAAA 4
 RESULT 15

BM971435/c
LOCUS 657 bp mRNA linear EST 21-MAR-2002
DEFINITION UI-CF-DUI-abd-1-16-0-UI.s2 UI-CF-DUI Homo sapiens cDNA clone
UI-CF-DUI-abd-1-16-0-UI 3', mRNA sequence.
ACCESSION BM971435
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Bonaldo M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: McCray, PB
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-30, >AT-richLow_complexity
Seq primer: M13 FORWARD
POLYA=Yes

FEATURES
source
1..657
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-abd-1-16-0-UI"
/clone_lib="UI-CF-DUI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into pT73-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (GT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC
BASE COUNT 155 a 114 c 150 g 237 t 1 others
ORIGIN

Query Match 14.1%; Score 454.6; DB 14; Length 657;
Best Local Similarity 82.5%; Pred. No. 2.9e-80;
Matches 520; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 917 GTCACTCTGCTAATTCACCCCTTCGCCACTACAGTCACTGCCATCACACCCGCTCCTA 976
|||||
DB 641 GTCACTCTGCTAATTCACCCCTTCGCCACTACAGTCACTGCCATCACACCCGCTCCTA 582
QY 977 TTCAACAGCCCGTAAAGCATCGGCTCTGAAACCCAGAGCTGCAGACTGCTTTAGCCC 1036
|||||

Db 581 TTCAACAGCCCTGTGAAAGTATGCGTGTATTAACACAGAGCTACAGACTGCTTTAGCAC 522
QY 1037 CAACCCATCCTTCTTGATGCCACAGCCAGTTCAGACTGTTCAGCCCTACCCCTTTTCTG 1096
|||||
Db 521 CTACACACCCCTTCTTGATGCCACAGCCCAATTCAAACTGTTCAACCCAGTCCCTTTCTG 462
QY 1097 AGGTACAGCTTCAAGTGTGCTGTATCCCACTCTGTTCAGCTGTTCGCTGAAGCTTCAAGTATCAAG 1156
|||||
Db 461 AGGAACCGCTTCAATGTGACTGTGATGCACCTGTTCGCTGAAGCTTCAAGTATCAAG 402
QY 1157 GTCCACACCCGCTTATCCAAACATCTGTACACCAAAACCCATCTGTCCCTCCATATG 1216
|||||
Db 401 GACACACCCACCCCTACCCAAACATCTGTGCACCAAAACCCATCTGTTCCTCCATACG 342
QY 1217 AGTCAGTAAGTAAGCCCTGCAAGAGATGAACAGCTTACCCCAAGAAAGATGATAGT 1276
|||||
Db 341 AGTCAATCACTAAGCCTAGCAAGAGAGATCAGCAAGCTTGCCCAAGAAAGATGAGAGT 282
QY 1277 AGAAGAGTGGGACAGTGGTGTACTCTGGGGATTAAGAAAGAAACAGATTCACACTTCAC 1336
|||||
Db 281 AAAAGAGTTATGAAATGTTGATAGTGGGGATTAAGAAAGAAACAGATTCACACTTCAC 222
QY 1337 CTATCACTGTTCGAAACCAAGAGATGAAGAAAGAGATCAGCAAGCTTGCCCAAGAAAGT 1396
|||||
Db 221 CTATCACTGTTCGAAACCAAGAGATGAAGAAAGAGATCAGCAAGCTTGCCCAAGAAAGT 162
QY 1397 ACTCCACACAGGCTTTAAGCTTCTTCATGGAGCAGCAGCTAGAGAGCTCTCTGAAAGTCTC 1456
|||||
Db 161 ATTCTCTCAAGCATTTAATTTCTTTATGGAGCAACATGTAGAAAATGTACTCAATCTC 102
QY 1457 ATCAGCAGCTGTGCATCGGAAGAGCAGCTAGAAAATGAAATGATCGGGGTGGATAT 1516
|||||
Db 101 ATCAGCAGCTGTGCATCGGAAGAGCAGCTAGAAAATGAAATGATCGGGGTGGATAT 42
QY 1517 CTCAAGATGCCAGGATCAATGAGAAAGA 1546
|||||
Db 41 TTAAGATGCTCCATTTTATATCTTAAAAA 12

Search completed: January 16, 2003, 21:48:49
Job time: 2795.39 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 5238.57 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagcaccgccgaagt.....aagagcattatttggggg 3155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pt.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_pat.*
- 23: em_ph.*
- 24: em_pl.*
- 25: em_ro.*
- 26: em_sts.*
- 27: em_un.*
- 28: em_vl.*
- 29: em_htg_hum.*
- 30: em_htg_inv.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3155	100.0	3155	6	AR201457	AR201457 Sequence
2	3097.4	98.2	3460	10	AB023958	AB023958 Mus muscu
3	1749.6	55.5	5486	6	E38226	E38226 Human tumor
4	1748	55.4	5486	6	E38227	E38227 Human tumor
5	1746.4	55.4	4098	9	AF207547	AF207547 Homo sapi
6	1717.2	54.4	3412	9	AB028019	AB028019 Homo sapi
7	812.6	25.8	3213	6	AR201456	AR201456 Sequence
8	812.6	25.8	3213	10	AF104414	AF104414 Mus muscu
9	760	24.1	3984	6	AR201455	AR201455 Sequence
10	760	24.1	4241	9	AF164041	AF164041 Homo sapi
11	760	24.1	4307	9	AF104413	AF104413 Homo sapi
12	747.8	23.7	2442	6	E24613	E24613 warts prote
13	742.8	23.5	57393	2	AC094779	AC094779 Rattus no
14	676.6	21.4	1374	6	E24614	E24614 warts prote
15	566.2	17.9	5360	3	DROWARTS	DROWARTS
16	564.6	17.9	5720	3	DMU296087	DMU296087
17	564.6	17.9	5720	6	AR201454	AR201454 Sequence
18	495.8	15.7	183597	9	AL356285	AL356285 Human DNA
19	460.4	14.6	74757	2	AC118930	AC118930 Mus muscu
20	366.6	11.6	16490	2	AC102712	AC102712 Mus muscu
21	356.8	11.3	181750	2	AC098623	AC098623 Rattus no
22	352.4	11.2	80146	9	AL583963	AL583963 Human DNA
23	311.6	9.9	335259	2	AC015834	AC015834 Homo sapi
24	287.2	9.1	2305	8	NCOT1	X97657 N crassa mR
25	286.2	9.1	2231	8	AF041843	AF041843 Ustilago
26	258.8	8.2	2160	6	AR170896	AR170896 Sequence
27	258.4	8.2	1745	3	AF275634	AF275634 Caenorhab
28	251	8.0	215989	2	AC102227	AC102227 Mus muscu
29	248.2	7.9	28133	2	AC012975	AC012975 Drosophil
30	248.2	7.9	163741	3	AC007821	AC007821 Drosophil
31	247.4	7.8	2126	3	AF238490	AF238490 Drosophil
32	243.6	7.7	1235	3	CEPROKINX	Z34989 C.elegans m
33	243.4	7.7	236109	3	AE003775	AE003775 Drosophil
34	241.8	7.7	2265	3	AY051880	AY051880 Drosophil
35	240	7.6	245067	2	AC087136	AC087136 Mus muscu
36	237.8	7.5	2100	3	DMPROKINX	Z35103 D.melanog
37	237.8	7.5	2114	3	AF239171	AF239171 Drosophil
38	237	7.5	4725	9	BC028603	BC028603 Homo sapi
39	237	7.5	4983	6	AX056364	AX056364 Sequence
40	237	7.5	5181	9	AB023182	AB023182 Homo sapi
41	233.8	7.4	1935	6	AR139101	AR139101 Sequence
42	231.6	7.3	1579	8	AY059113	AY059113 Arabidops
43	231.6	7.3	1894	8	AY086014	AY086014 Arabidops
44	231.6	7.3	1970	8	AY034932	AY034932 Arabidops
45	231.4	7.3	2422	8	AY128279	AY128279 Arabidops

ALIGNMENTS

RESULT 1	AR201457	AR201457	3155 bp	DNA	linear	PAT 20-APR-2002
LOCUS	Sequence 7 from patent US 6359193.					
DEFINITION	AR201457					
ACCESSION	AR201457.1	GI:20252345				
VERSION	Unknown.					
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 3155)					
AUTHORS	Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.					
TITLE	Nucleotide sequences of lats genes					
JOURNAL	Patent: US 6359193-A 7 19-MAR-2002;					
FEATURES	Location/Qualifiers					

Qy	961	AGGACCAACTCTCTTCAACAACCCAAACACTGAGCGCCTCACTGCGCGGCCCCCAACACGGTTC	1020
Db	961		1020
Qy	961	AGGACCAACTCTCTTCAACAACCCAAACACTGAGCGCCTCACTGCGCGGCCCCCAACACGGTTC	1020
Db	961		1020
Qy	1021	ACCGCGTGACGCGCCGACACACTCTTCAACCGCTGTGAAGACGCTGCGTGTGCTGCGGCGCC	1080
Db	1021		1080
Qy	1021	ACCGCGTGACGCGCGGACACATCTTCAACCGCTGTGAAGACGCTGCGTGTGCTGCGGCGCC	1080
Db	1021		1080
Qy	1081	GAGCCCCAGACAGCCGTGGGGCCCTCGCACCCCGCCTGGGTGGGTGGCGCCCAACAGCACCT	1140
Db	1081		1140
Qy	1081	GAGCCCCAGACAGCCGTGGGGCCCTCGCACCCCGCCTGGGTGGGTGGCGCCCAACAGCACCT	1140
Db	1081		1200
Qy	1141	GCCACTGAGAGCCTGGAGACGAAGGAGGCGACGCGAGCGCCACACCCGCTGGATGTGGAC	1200
Db	1141		1200
Qy	1141	GCCACTGAGAGCCTGGAGACGAAGGAGGCGACGCGAGCGCCACACCCGCTGGATGTGGAC	1200
Db	1141		1200
Qy	1201	TATGGCGCTCCGAGCGCAGGTGCCACACCGCTCCGTATCCAAAGACATTGCTGCTGCC	1260
Db	1201		1260
Qy	1201	TATGGCGCTCCGAGCGCAGGTGCCACACCGCTCCGTATCCAAAGACATTGCTGCTGCC	1260
Db	1201		1320
Qy	1261	AGTAAGTCTGAGCAGTACAGCTGGACAGCTGTGCACCAAGTGTCCACAGACAGT	1320
Db	1261		1320
Qy	1261	AGTAAGTCTGAGCAGTACAGCTGGACAGCTGTGCACCAAGTGTCCACAGACAGT	1320
Db	1261		1380
Qy	1321	CTGCGAGGGGCACTGATCTAGACGGAGTGTACAGAGCGCACAAAGTGTCCGAAGGGAGAC	1380
Db	1321		1380
Qy	1321	CTGCGAGGGGCACTGATCTAGACGGAGTGTACAGAGCGCACAAAGTGTCCGAAGGGAGAC	1380
Db	1321		1440
Qy	1381	AAAGCTGGCAGAGACAAAAAGCAGATTACAGACCTCCCGGTGCCTGTCCGCAAGAATAGC	1440
Db	1381		1440
Qy	1381	AAAGCTGGCAGAGACAAAAAGCAGATTACAGACCTCCCGGTGCCTGTCCGCAAGAATAGC	1440
Db	1381		1500
Qy	1441	AGAGATGAAGAGAGACAGAGTCTCCCATCAAGAGTTACTCCCTTATGCCTTCAAAATTC	1500
Db	1441		1500
Qy	1441	AGAGATGAAGAGAGACAGAGTCTCCCATCAAGAGTTACTCCCTTATGCCTTCAAAATTC	1500
Db	1441		1560
Qy	1501	TTCATTGAGGACAAACGCTGGAGAATGTCAACAAACCTACCAGCAGAAGGTGACGGGAGG	1560
Db	1501		1560
Qy	1501	TTCATTGAGGACAAACGCTGGAGAATGTCAACAAACCTACCAGCAGAAGGTGACGGGAGG	1560
Db	1501		1620
Qy	1561	CTACAGCTGGAGCAGGAATGGCCAAAGCTGGGCTCTGTGAGCGCGACGAGGACAGATG	1620
Db	1561		1620
Qy	1561	CTACAGCTGGAGCAGGAATGGCCAAAGCTGGGCTCTGTGAGCGCGACGAGGACAGATG	1620
Db	1561		1680
Qy	1621	AGGAAGATCCTCTACCAAGAAGGACTTAACATAACACCGGCTGAAGAGGCCAAGATGAC	1680
Db	1621		1680
Qy	1621	AGGAAGATCCTCTACCAAGAAGGACTTAACATAACACCGGCTGAAGAGGCCAAGATGAC	1680
Db	1621		1740
Qy	1681	AGTCCATGTTGTGAANAATCAAGACTTGAAGCATCGTCTTTGGGGAAGTGTGCCTC	1740
Db	1681		1740
Qy	1681	AGTCCATGTTGTGAANAATCAAGACTTGAAGCATCGTCTTTGGGGAAGTGTGCCTC	1740
Db	1681		1800
Qy	1741	GCTTTAAGCTGGACACTACGCTCTGTAGCCCATGAAGACTCTCAGGAAGAAGGATGTC	1800
Db	1741		1800
Qy	1741	GCTTTAAGCTGGACACTACGCTCTGTAGCCCATGAAGACTCTCAGGAAGAAGGATGTC	1800
Db	1741		1860
Qy	1801	CTGAACCCGAATCAAGTGGCCCATGTCAAGCTGTAGAGGACATCTGCTGCTGAACGACAC	1860
Db	1801		1860
Qy	1801	CTGAACCCGAATCAAGTGGCCCATGTCAAGCTGTAGAGGACATCTGCTGCTGAACGACAC	1860
Db	1801		1920
Qy	1861	AATGATGGGTGGTCAAACTCTACTCTCTCCAGGACAAGGACAGCCTGACTTTGTG	1920
Db	1861		1920
Qy	1861	AATGATGGGTGGTCAAACTCTACTCTCTCCAGGACAAGGACAGCCTGACTTTGTG	1920
Db	1861		1980
Qy	1921	ATGGACTACATACCAAGCGGGGATATGATGAGCCTGCTGATCAGGATGAGGCTTCCCT	1980
Db	1921		1980
Qy	1921	ATGGACTACATACCAAGCGGGGATATGATGAGCCTGCTGATCAGGATGAGGCTTCCCT	1980
Db	1921		204
Qy	1981	GAGCACTTGGCCGCTTCTACATTTGCAGAGTTGACCTGGCCATTCAAAGTGTCCACAAG	204
Db	1981		204
Qy	1981	GAGCACTTGGCCGCTTCTACATTTGCAGAGTTGACCTGGCCATTCAAAGTGTCCACAAG	204
Db	1981		2100
Qy	2041	ATGGGCTTTATCCACCGGGACATCAAGCCTGACAAATCTCATGCACTCATGCACTGGTCTAT	2100

```

Db 2041 ATGGGCTTTATCCAGCGGACATCAAGCTGACAACTACTCATCGACCTGGATGTCAT 2100
QY 2101 ATTAAGCTACAGATTTTGGCTCTGCACCTGATTCAGTGGACTCACAATTTCCAGTAG 2160
Db 2101 ATTAAGCTACAGATTTTGGCTCTGCACCTGATTCAGTGGACTCACAATTTCCAGTAG 2160
QY 2161 TACCAGAAAGGAACACATGAGACGAGCAGATGAGCCCGGTGACCTCTGGGACGAT 2220
Db 2161 TACCAGAAAGGAACACATGAGACGAGCAGATGAGCCCGGTGACCTCTGGGACGAT 2220
QY 2221 GTTTCACACTGCTGTGGAGACAGTTAAAGACCCCTGAGCAGAGCGCGCAGAGCAG 2280
Db 2221 GTTTCACACTGCTGTGGAGACAGTTAAAGACCCCTGAGCAGAGCGCGCAGAGCAG 2280
QY 2281 CACGAGGTGCTGCGCACATTTCTTGTGCGGACACCAATATACATCCTCGGAGGTG 2340
Db 2281 CACGAGGTGCTGCGCACATTTCTTGTGCGGACACCAATATACATCCTCGGAGGTG 2340
QY 2341 CTTCTCCGCAAGGGTACACGAGCTCTGCTGCTGAGCGTGGTGTGATTTCTCTTT 2400
Db 2341 CTTCTCCGCAAGGGTACACGAGCTCTGCTGCTGAGCGTGGTGTGATTTCTCTTT 2400
QY 2401 GAGATGCTGGTGGGAGCGCGCTTTCTTGTGGCCCCACCCACACAGACGAGCTGAAG 2460
Db 2401 GAGATGCTGGTGGGAGCGCGCTTTCTTGTGGCCCCACCCACACAGACGAGCTGAAG 2460
QY 2461 GTGATCACTGGGAGACGCTGCATATCCCTACGACGAGTGAAGCTCAGCGCTGAGGCC 2520
Db 2461 GTGATCACTGGGAGACGCTGCATATCCCTACGACGAGTGAAGCTCAGCGCTGAGGCC 2520
QY 2521 CGAGACCTCATCAGCAAGCTGTGCTGCGCGGCTGACTGCGCGCTGGCAGGATGGGCA 2580
Db 2521 CGAGACCTCATCAGCAAGCTGTGCTGCGCGGCTGACTGCGCGCTGGCAGGATGGGCA 2580
QY 2581 GATGACCTCAAGGACACCCGCTTTCTTCAACACCATCGACTTTTCCCGTGACATCCGAAG 2640
Db 2581 GATGACCTCAAGGACACCCGCTTTCTTCAACACCATCGACTTTTCCCGTGACATCCGAAG 2640
QY 2641 CAGGCTGACACCTAGCTCCCGACCATCAGCCACCCATGAGACCTCCAAATTTGACCGG 2700
Db 2641 CAGGCTGACACCTAGCTCCCGACCATCAGCCACCCATGAGACCTCCAAATTTGACCGG 2700
QY 2701 GTGGATGAAGAAAGCCCTTGGCAGAGCGGACGAGCAGCGCCGAGGAGCTGGGACAG 2760
Db 2701 GTGGATGAAGAAAGCCCTTGGCAGAGCGGACGAGCAGCGCCGAGGAGCTGGGACAG 2760
QY 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTCACCTTCCGAG 2820
Db 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTCACCTTCCGAG 2820
QY 2821 TTTCTTGATGACAGCGCTATCCCTTCCGCTGCGCGAAGCCCTCAGAGCCCGCAGAGT 2880
Db 2821 TTTCTTGATGACAGCGCTATCCCTTCCGCTGCGCGAAGCCCTCAGAGCCCGCAGAGT 2880
QY 2881 GCAGACCCAGGGGATGCGGACTTGAAGGTGCGCGCGGAGGCTGCGAGCGGCTGACGTTG 2940
Db 2881 GCAGACCCAGGGGATGCGGACTTGAAGGTGCGCGCGGAGGCTGCGAGCGGCTGACGTTG 2940
QY 2941 TAAGCCTCAGTTAACCACAACTCGAGGAACCCAAATAGATTTCTTTTTCAGAGACAA 3000
Db 2941 TAAGCCTCAGTTAACCACAACTCGAGGAACCCAAATAGATTTCTTTTTCAGAGACAA 3000
QY 3001 ACTCAAGCTTAGGATCCTTCAATTTTGTGCTGTAATGCGCAACGAGAGCTCAAC 3060
Db 3001 ACTCAAGCTTAGGATCCTTCAATTTTGTGCTGTAATGCGCAACGAGAGCTCAAC 3060
QY 3061 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAAA 3120
Db 3061 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAAA 3120
QY 3121 TTAGTACAGTGAAGAGACACTTATTTTTGGGG 3155

```

```

Db 3121 TTAGTACAGTGAAGAGACACTTATTTTTGGGG 3155

RESULT 2
LOCUS AB023958
DEFINITION Mus musculus MmLATS2 mRNA for warts/lats-like kinase, complete cds.
ACCESSION AB023958
VERSION AB023958.1 GI:7212785
KEYWORDS MmLATS2; warts/lats-like kinase;
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Yabuta,N., Fujii,T., Copeland,N.G., Gilbert,D.J., Jenkins,N.A., Nishiguchi,H., Endo,Y., Toji,S., Tanaka,H., Nishimune,Y. and Nojima,H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
AUTHORS 2 (bases 1 to 3460)
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for Microbial Diseases, Osaka University, Department of Molecular Genetics; Yamadaoka 3-1, Suita, Osaka 565-0871, Japan (E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980, Fax:81-6-6875-5192)
COMMENT Sequence updated (06-Jan-2000).
FEATURES
Source 1..3460
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..3460
/gene="MmLATS2"
CDS 134..3262
/gene="MmLATS2"
/note="WALK"
/codon_start=1
/product="warts/lats-like kinase"
/db_xref="GI:7212786"
/translation="MRPTFTATTYSGNRRLQREIRGLKQPSKASTQGLLVGNPSD
TSLDAKVLGSKDASRQOQMRATPKPGYKALREIRYSLLPFANESYSAAEVNRQM
LQELVNAAGCDQEMAGRALKQTSRSIEAALYISKMYLDPNREIIVRIQTSFGKG
LAPTPTVRPSEGTGEALPSYHRLGANYEGPAALEMPQYLDLFLPGAGAGTHGA
QAHQHPKGYSTAVEPSAHFPCTHYGRGHLLSEOPGYQVRSSEFONKTPDAYSSMA
KAQGGPASPALTFAHAGLYTASHHKPAATPGAHPLHVLGTGPTFTGESSAQAVLAP
SRNSLNADLYELGSTVPSAAPLARRDSLOKGLEASRPHVAFRAGPSRTSNPNQP
EPSPAPNTVTVAATHLHPKSVRLPEPTAVGSPHAWAAPATAPATESLETK
EGSAGHPHLDVYDGGSKRRCPYPYKHLHLLSEQYSVDLSLCTSVQOQSLRGTE
ODRDKSHKAGKAGKDRKKQIQTSPVVRKNSRDEKRESIKSYSPAFKTFMEQ
HVENVITYQOKVSRRLQLEQAVAKCEAQEMRKVLKQVNSYKRLKRAKMDKS
MFVKIKTLGIGAFGEVCLACKDLTHALVAKETLRKQDLVLRNOVAHVKAERDILAEAD
NEWVKLYSFQDKSLYFYVDYIFPGDMMSLLIRMEVPEHLARFYIAETLAIESV
HKMGFIHRDIKPNILDLIDGLCTGFRTWTHNSKYQGNHMRDSDMEPGD
LMDVDSNCRCCDRKLTLEQRAQKORCLASHLVGTPTNYIAPEVLLRKYTOLCDDWS
VGVLTFEMLVGQPPFLAPTPTTQTKVINWSTLHIPTQVRLSAEARDLITKLCDAAD
CRLGDGADDLKAHPFNTIDFSDIRKQAPAPYPTISHPMDTINFDVDESPWHEA
SGSAKAWDTLASPSKHPHAHYEFTFRFEDDNGYFFRCRKPSEPAESADPADL
EGAAEGCOPVIV"
polyA_site 3440
/gene="MmLATS2"
BASE COUNT 861 a 992 c 963 g 644 t
ORIGIN
Query Match 98.2%; Score 3097.4; DB 10; Length 3460;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3113; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAGAGCCACCCCGAAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 60

```

Db 320 ATGAGACCAACCCGAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATCC 379
 QY 61 CTCTGCTTTTCCACAGAGTTCAGGACATTCGGAGCTGCAGAGGTGAACCGGAGATG 120
 Db 380 CTCTGCTTTTCCACAGAGTTCAGGACATTCGGAGCTGCAGAGGTGAACCGGAGATG 439
 QY 121 CTTGAGAGTTGGTGAATCGGAGTGTACACAGAGATGCTGCAGAGGCTCAGCGAG 180
 Db 440 CTTGAGAGTTGGTGAATCGGAGTGTACACAGAGATGCTGCAGAGGCTCAGCGAG 499
 QY 181 ACGGCAAGTAGGAGTATCGAAGTTCGCTTGGAGTACATCAGTAAAGTGGCTACCTGGAC 240
 Db 500 ACGGCAAGTAGGAGTATCGAAGTTCGCTTGGAGTACATCAGTAAAGTGGCTACCTGGAC 559
 QY 241 CCCAGGAATGAGCAGATTTGCGAGTGTATCAAGCAGACCTCCCGAGGAAGGCGCTGGCG 300
 Db 560 CCCAGGAATGAGCAGATTTGCGAGTGTATCAAGCAGACCTCCCGAGGAAGGCGCTGGCG 619
 QY 301 TCCACCCGCTGACTCGCGGCCAGTTTCGAGGGCACAGGGGAAGCAGCTCCCATCTAC 360
 Db 620 CCCACCCGCTGACTCGCGGCCAGTTTCGAGGGCACAGGGGAAGCAGCTCCCATCTAC 679
 QY 361 CACGAGCTGGTGTGCAAACTACGAGGGCCCGCCGCACTGGAGGAGATGCCCGGCA 420
 Db 680 CACGAGCTGGTGTGCAAACTACGAGGGCCCGCCGCACTGGAGGAGATGCCCGGCA 739
 QY 421 TATTAGACTTTCCTTCCCTGGAGCGGAGCGGACCCAGCTGCGCAGGCTCACAG 480
 Db 740 TATTAGACTTTCCTTCCCTGGAGCGGAGCGGACCCAGCTGCGCAGGCTCACAG 799
 QY 481 CATCTCCAAAGGTTACAGCAGCAGTATAGACCAAGTGGCACTTTCGGGGCACAC 540
 Db 800 CATCTCCAAAGGTTACAGCAGCAGTATAGACCAAGTGGCACTTTCGGGGCACAC 859
 QY 541 TATGGTGTGTCTACTACTATCGGAGCTGCTGGGTATGGGGTGCAGCGAGTTCCTCC 600
 Db 860 TATGGTGTGTCTACTACTATCGGAGCTGCTGGGTATGGGGTGCAGCGAGTTCCTCC 919
 QY 601 TTCAGAACAGACCCACAGATGCTTATTCAGAGTGGCAAGGCGGAGGCTGGCGCT 660
 Db 920 TTCAGAACAGACCCACAGATGCTTATTCAGAGTGGCAAGGCGGAGGCTGGCGCT 979
 QY 661 CCGCAGCCTCCTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 980 CCGCAGCCTCCTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1039
 QY 721 CGGGTACCCACCTGGGGCCACCCATTATGTTGGGACCCGCGGCTCCACGTTT 780
 Db 1040 CGGGTACCCACCTGGGGCCACCCATTATGTTGGGACCCGCGGCTCCACGTTT 1099
 QY 781 ACTGGCAAGCTCTGCACAGCTGTGCTGCAAGCTGCTGCAAGCAGCTCAATGCTGAC 840
 Db 1100 ACTGGCAAGCTCTGCACAGCTGTGCTGCAAGCTGCTGCAAGCAGCTCAATGCTGAC 1159
 QY 841 TTGTACAGCTGGGCTCCACGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 Db 1160 TTGTACAGCTGGGCTCCACGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
 QY 901 CTGCAGAGCAGGCTTGAAGCCTTCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 Db 1220 CTGCAGAGCAGGCTTGAAGCCTTCGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1279
 QY 961 AGGACCAACTCTTCAACAACCCACAGCTGAGCCCTCAGTCCCGGCGGCGGCGGCTGCTGCT 1020
 Db 1280 AGGACCAACTCTTCAACAACCCACAGCTGAGCCCTCAGTCCCGGCGGCGGCGGCTGCTGCT 1339
 QY 1021 ACCGCGGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 Db 1340 ACCGCGGTGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1399
 QY 1081 GAGCCCCAGACAGCGGTGGGGCGGCTGCGACCCGCGGCTGGGCTGGGCGGCGGCGGCGGCGG 1140

Db 1400 GAGCCCCAGACAGCCGCTGGGGCCCTCGCACCCCGCTGGGTGGCTGCCCGCCACAGCACT 1459
 QY 1141 GCCACTGAGAGCTTGAGACGAAGGAGGCGAGCGAGGCGCCACACCGCTGGATGTGGAC 1200
 Db 1460 GCCACTGAGAGCTTGAGACGAAGGAGGCGAGCGAGGCGCCACACCGCTGGATGTGGAC 1519
 QY 1201 TATGGCGGCTCCGAGCGGAGTGCACCGCTCCGCTATCCAAAGACATTCGTGCTGCC 1260
 Db 1520 TATGGCGGCTCCGAGCGGAGTGCACCGCTCCGCTATCCAAAGACATTCGTGCTGCC 1579
 QY 1261 AGTAAGCTTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCACAGTGTGCAGAGAGT 1320
 Db 1580 AGTAAGCTTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCACAGTGTGCAGAGAGT 1639
 QY 1321 CTGCGAGGGGCACTCATCTAGACGGGAGTGCACAGAGCAGCAAGAGTGCAGAGGAGAC 1380
 Db 1640 CTGCGAGGGGCACTCATCTAGACGGGAGTGCACAGAGCAGCAAGAGTGCAGAGGAGAC 1699
 QY 1381 AAAGCTGGCAGAGACAAAAGCAGATTCAGACCTCCCGGCTGCTGCCAGAGATAGC 1440
 Db 1700 AAAGCTGGCAGAGACAAAAGCAGATTCAGACCTCCCGGCTGCTGCCAGAGATAGC 1759
 QY 1441 AGAGATCAAGAGAGAGAGTCTCGCATCAAGAGTTCCTCCCTTATGCCCTTCAAAATTC 1500
 Db 1760 AGAGATCAAGAGAGAGAGTCTCGCATCAAGAGTTCCTCCCTTATGCCCTTCAAAATTC 1819
 QY 1501 TTGATGGAGCAACACGCTGGAGATGTCTATCAAAACCTACAGCAGAGGTGACCGGAGG 1560
 Db 1820 TTGATGGAGCAACACGCTGGAGATGTCTATCAAAACCTACAGCAGAGGTGACCGGAGG 1879
 QY 1561 CTACAGCTGGAGCAGGAATGGCCAAAGCTGGGCTCTGTGAGGCGCGAGGAGAGATG 1620
 Db 1880 CTACAGCTGGAGCAGGAATGGCCAAAGCTGGGCTCTGTGAGGCGCGAGGAGAGATG 1939
 QY 1621 AGGAAGTCTCTACCAAGAGAGTCTTAACTACAAACCGCTGAAGAGGCGCAAGATGGAC 1680
 Db 1940 AGGAAGTCTCTACCAAGAGAGTCTTAACTACAAACCGCTGAAGAGGCGCAAGATGGAC 1999
 QY 1681 AAGTCCATGTTGTGAAAATCAAGACTCTAGGATCGGCTGCTGGGGAAGTGTGCCCT 1740
 Db 2000 AAGTCCATGTTGTGAAAATCAAGACTCTAGGATCGGCTGCTGGGGAAGTGTGCCCT 2059
 QY 1741 GCTTGAAGCTGGACACTCAGCTCTAGCCATGAAGACTCTCAGGAAGAGAGATGTC 1800
 Db 2060 GCTTGAAGCTGGACACTCAGCTCTAGCCATGAAGACTCTCAGGAAGAGAGATGTC 2119
 QY 1801 CTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCGGCTGAGCAGAC 1860
 Db 2120 CTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCGGCTGAGCAGAC 2179
 QY 1861 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGACAGCTGTACTTTGTG 1920
 Db 2180 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGACAGCTGTACTTTGTG 2239
 QY 1921 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGACAGCTGTACTTTGTG 1980
 Db 2240 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGACAGCTGTACTTTGTG 2299
 QY 1981 GAGCAGCTGGCGGCTTCTACATTCAGAGTGTGACCTGGCCATTTGAAGTGTCCCAAG 2040
 Db 2300 GAGCAGCTGGCGGCTTCTACATTCAGAGTGTGACCTGGCCATTTGAAGTGTCCCAAG 2359
 QY 2041 ATGGGCTTATCCACCGGACATCAAGCTTCCAGCAGAGGACAGCTGTACTTTGTG 2100
 Db 2360 ATGGGCTTATCCACCGGACATCAAGCTTCCAGCAGAGGACAGCTGTACTTTGTG 2419
 QY 2101 ATTAAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGAGTCAAAATTCAGATAC 2160
 Db 2420 ATTAAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGAGTCAAAATTCAGATAC 2479
 QY 2161 TACCAGAAAGGAAACCATGAGCAGAGCAGATGAGCGGCGGCTGACCTCTGGGAGAT 2220
 Db 2480 TACCAGAAAGGAAACCATGAGCAGAGCAGATGAGCGGCGGCTGACCTCTGGGAGAT 2539

QY 2221 GTTTCACCTGCTGCTGGAGACAGTTAAAGACCTTGAGCAGAGGCGCAGAAAGCAG 2280
 Db 2540 GTTTCACCTGCTGCTGGAGACAGTTAAAGACCTTGAGCAGAGGCGCAGAAAGCAG 2599
 QY 2281 CACCAGAGTGCCTGCACATCTCTGTCGGGACACCAATATACATCGCTCGGAGGTG 2340
 Db 2600 CACCAGAGTGCCTGCACATCTCTGTCGGGACACCAATATACATCGCTCGGAGGTG 2659
 QY 2341 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGCTGCGTGTGATCTCTTT 2400
 Db 2660 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGCTGCGTGTGATCTCTTT 2719
 QY 2401 GAGATCTGTTGGGACGCGCTTCTTGGCCCCCACCACCCACAGAGCAGCTGAAG 2460
 Db 2720 GAGATCTGTTGGGACGCGCTTCTTGGCCCCCACCACCCACAGAGCAGCTGAAG 2779
 QY 2461 GTGATCACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTAGCGCTGAGGCC 2520
 Db 2780 GTGATCACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTAGCGCTGAGGCC 2839
 QY 2521 CGAGACCTCATCAGCAAGCTGTGCTGCGGCTGACTGCGCTGGCAGGATGGGCA 2580
 Db 2840 CGAGACCTCATCAGCAAGCTGTGCTGCGGCTGACTGCGCTGGCAGGATGGGCA 2899
 QY 2581 GATGACCTCAAGCAGCACCCTTCTTCAACACCATCGACTTTTCCGCTGACATCCGAAAG 2640
 Db 2900 GATGACCTCAAGCAGCACCCTTCTTCAACACCATCGACTTTTCCGCTGACATCCGAAAG 2959
 QY 2641 CAGGCTGACACCTTACCTCCGACCAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2700
 Db 2960 CAGGCTGACACCTTACCTCCGACCAATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3019
 QY 2701 GTGGATGAGAAAGCCCTTGGCAGAGGCGCAGGAGAGAGGCGCAGGAGGCGGACAG 2760
 Db 3020 GTGGATGAGAAAGCCCTTGGCAGAGGCGCAGGAGAGAGGCGCAGGAGGCGGACAG 3079
 QY 2761 CTGGCTCCGACAGCAAGCATCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2820
 Db 3080 CTGGCTCCGACAGCAAGCATCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3139
 QY 2821 TTCTTCGATGACAAAGCTATCCCTTCCGCTGCGGCGGAGGCTGCGGCGGCTGACGTG 2880
 Db 3140 TTCTTCGATGACAAAGCTATCCCTTCCGCTGCGGCGGAGGCTGCGGCGGCTGACGTG 3199
 QY 2881 GCAGACCCAGGGATGCGGACTTGGAGCTGCGGCGGAGGCTGCGGCGGCTGACGTG 2940
 Db 3200 GCAGACCCAGGGATGCGGACTTGGAGCTGCGGCGGAGGCTGCGGCGGCTGACGTG 3259
 QY 2941 TAAGCCTCAGTTAAACCACTCGAGAAACCCAAATGAGATTTCTTTTCAGAAACAA 3000
 Db 3260 TAAGCCTCAGTTAAACCACTCGAGAAACCCAAATGAGATTTCTTTTCAGAAACAA 3319
 QY 3001 ACTCAAGCTTAGAATCCCTTCAATTTTCTGTTAAATGGCAACAGAGAGTCAAC 3060
 Db 3320 ACTCAAGCTTAGAATCCCTTCAATTTTCTGTTAAATGGCAACAGAGAGTCAAC 3379
 QY 3061 ATGATTTCAATTAAGCCTCTGAGGACCTCACTGATTAACAGATTTTAAATAA 3120
 Db 3380 ATGATTTCAATTAAGCCTCTGAGGACCTCACTGATTAACAGATTTTAAATAA 3439
 QY 3121 TTAGTACAGTATGGAAGA 3139
 Db 3440 TAAAAAATAAAAAA 3458

RESULT 3
 E38226
 LOCUS E38226 5486 bp DNA linear PAT 31-JAN-2002
 DEFINITION Human tumor regulatory gene.
 ACCESSION E38226
 VERSION E38226.1 GI:18626935
 KEYWORDS JP 2000210086-A/1.

SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo;
 REFERENCE 1 (bases 1 to 5486)
 AUTHORS Koga, J., Kono, K. and N.Z.F.
 TITLE Human tumor regulatory gene
 JOURNAL Patent: JP 2000210086-A 1 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 COMMENT OS Homo sapiens (human)
 PN JP 2000210086-A/1
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
 C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68, PC
 G01N33/50,
 PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09, PC
 C12R1:91),C12N15/00,
 PC A61K37/36,(C12N15/00,C12R1:91)
 CC
 FH Key Location/Qualifiers
 FT source 1..5486
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 1378 a 1452 c 1390 g 1266 t
 ORIGIN
 Query Match 55.5%; Score 1749.6; DB 6; Length 5486;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 2389; Conservative 0; Mismatches 544; Indels 153; Gaps 16;
 QY 1 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 60
 Db 576 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 635
 QY 61 CTCCTGCTTTTGGCCACGAGTCAGGCACCTTCGGCAGCTGCGAGGTGAACCGGAGATG 120
 Db 636 TTGTTGCTTTTGGCTTAATGAATCGGCGACCTCTGCAGCTGCGAGGTGAACCGGAAATG 695
 QY 121 CTTGAGGAGTGTGTAATCGCGCATGTGACGAGGAGATGGTGGCAGAGCGCTCAGCAG 180
 Db 696 CTGAGGAACTGTGTAATCGCGCATGTGACGAGGAGATGGTGGCAGAGCGCTCAGCAG 755
 QY 181 ACGGCAGTAGGAGTATCGAAGCTGCTTTGGAGTACATCAGTAAGATGGGCTACCTGGAC 240
 Db 756 ACTGGCAGCAGGAGCATCGAGCGGCTGGAGTACATCAGCAAGATGGGCTACCTGGAC 815
 QY 241 CCAGGAATGAGCAGATTTGTGGAGTCAATCAGCAGACCTCCCGAGGAAGGCGCTGGCG 300
 Db 816 CCAGGAATGAGCAGATTTGTGGAGTCAATCAGCAGACCTCCCGAGGAAGGCGCTCATG 875
 QY 301 TCACCCCGGTCACCTCGGCGGCCAGTTTCGAGGGCAGCAGGGAAGCACTCCCATCTCTAC 360
 Db 876 CCAACCCAGTACGCGGAGGCGCCAGCTTCGAGGAGGAGGCGGATTCGTTGGTCTCTAC 935
 QY 361 CACCACTGGGTGGTGCAGAACTACGAG-----GGCCCCCGCCGA 399
 Db 936 CACCACTGGGTGGTGCAGAACTACGAG-----GGCCCCCGCCGA 995
 QY 400 CTGGAGGAGATGCGCGGCGCAATATTAGACTTTCTTCCCTGGAGCGCGGAGCGGACAC 459
 Db 996 CTGGAGGAGATGCGCGGCGCGTACGTTGGACTACCTTTTCCCGGAGTGGCGCCACACGG 1055
 QY 460 CACGTTGCCAGGCTCACCAGCATCTCCCAAGGTTACAGCAGCAGCAGTAGAGCAAGT 519
 Db 1056 CCGGGCCACAG---CACCAGCAGCCACCCAGGCGCTACGTTGCCAGCGCAGGAGGAGCA 1112
 QY 520 GCG-----CACTTTCCGGGCGACACACTATGTTGCTGTTGCTATCTACT-----ATCG 564

Db 1113 GGGGACACTTCCGCTGACAGGCGCGCACTACGCGGCGCGCACTGCTGTGCTGGG 1172
QY 565 GAGCAGTCTGGGTATGGGTGACGCGAGTTCCTCTTCCAGAACAAAGACGCGCACAGAT 624
Db 1173 GAACCCCTGGGTACGAGGTGACGCGAGCGCCCTCTTCCAGAGCAAGCGCGCGGAG 1232
QY 625 GCC-----TATTCACAGATGGC-----AAGGCCAGGGTGGGCCCTCC-----CGCCAGCCCTC 672
Db 1233 ACCGGGGGTATGCGCCAGCCTGCCACGAAGAGCGAGGAGGAGCGCGAGCGCGGCCCTC 1292
QY 673 ACCTTTCTCTGCCCATCTGGGTGTACACTGCTGCTGCACCAACAAGCGCGCGCTACCCCA 732
Db 1293 GCTTTCCACCCCTGCCCGCGGCTCTACGTGCGCCACCCACCAACAAGCAGCGCGGT 1352
QY 733 CTGGGGCCACCCATATACATGTGTGGGACACCGGGTCCACAGTTCCTTACTGGCGAAAGC 792
Db 1353 CCGCGGCGCCACAGCTGATGTGTGGCTCCCGCAGCCAGGTGTTCGCCAGCGACAGC 1412
QY 793 TCTGCAACAGCTGTGCTGCGACCGTCCAGGAACAGCTCAATGCTGACTGTGTACAGCTG 852
Db 1413 CCGCGCAGAGCTGTCTACTCTCCCTCGGGGAACAGCTCAACGTGGAGCTGTATGAATG 1472
QY 853 GG-----CTCCACAGGTGCGCTGTGCTGCTGAGCTCCACTGGCAGCGCGGACTCGCTGCG 906
Db 1473 GCGAGCAGCTCTCGTCCAGCAGTGGCGGCTGCCACCTTGGCGCGCGGAGCTCCCTGCG 1532
QY 907 AAGCAGGTCTAAGACCTTCGGGCG-----CGCATGTGGTTCGGGGCTGGC----- 954
Db 1533 AAGCGGGGCTGGAGCGCGCGCGCGCGAGCTGGCTTCGGGCTGACTGCCAGTG 1592
QY 955 CCACAGCAGCAACTCTCTCAACAAACCAAC-----CT 990
Db 1593 CCGCAGAGCACTCTTCAACAGCCACAGCGCGCGCGCTGCCCTGGCAAGGCC 1652
QY 991 GAGCCCTCACTGCCCGCCCAACAGCTCACCGCGGTGACGGCGGACACATCTTCCAC 1050
Db 1653 GAGCCCTCCCTGCCCGCCCAACAGCTGACGGCTGTACGGCGCGGACATCTTGAC 1712
QY 1051 CCGTGAAGAGGTGCTGCTGGGCGCGAGCGCCAGCGCGGCGGCGCTGCGAC 1110
Db 1713 CCGTGAAGAGGTGCTGCTGAGGCGCGGCGGCGGAGCGGTGTGGGCGCTCGAC 1772
QY 1111 CCGCCTGGGTGGTGGC-----CCACAGCAGCTGCGACATGAG 1149
Db 1773 CCGCCTGGGTGGCGGCGCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGAG 1832
QY 1150 AGCCTGGAGACGAGGA-----GGGAGCGAGCGCGCCACACCCCGCTG 1191
Db 1833 GCGTGGAGCGCAAGGAGGAGCATGCGCTGGCGGTGGCGCGGCGCGCTTCCCGCTG 1892
QY 1192 GATGTGAGTATGCGGCTCCGAGCGGAGTGGCCCGCTCGTATCCAAGCACTTG 1251
Db 1893 GACGTGAGTACGAGGCGCGGAGCGGAGGTGGCGCGCTCGCGCTACCGCGAGACCTG 1952
QY 1252 CTGCTGCCAGTAAGTCTGAGCAGTACAGGTGGACCTGGACAGCGCTGTGCACAGTGTG 1311
Db 1953 CTGCTGCGCAGCAAGTCSGAGCAGTAC-----GACCTGACAGCGCTGTGCGGAGGATG 2006
QY 1312 CAGCAGAGTCTGCGAGGGG-----CACTGTATCTAGCGGGAGTGACAGGCCACAAAGGT 1368
Db 2007 GAGCAGAGCTCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2066
QY 1369 GCGAAGGAGCAAGCTGCGCAGACAAAGAGATTCAGACCTCCCGGCTGCTGTC 1428
Db 2067 GCGAAGGAGCAAGCGCGGAAGGATTAAGAGCATTCAGACCTCTCCCGTTCCCGTC 2126
QY 1429 CGCAAGATAGCAGATGAGAGAGAGAGTCTCGCATCAAGAGTACTCCCTCTAT 1488
Db 2127 CGCAAAACAGCAGAGAGAGAGAGAGAGTCAAGCATCAAGAGTACTCGCATAC 2186
QY 1489 GCCTTCAATCTTCATGGAGCAACAGGTGGAGAAATGTCAACAAACCTTACCAGCAGAAG 1548
Db 2187 GCCTTAAAGTTCTTCATGGAGCAGCAGTGGAGATGTCAACAAACCTTACCAGCAGAAG 2246

QY 1549 GTACGCGGAGGCTACAGCTGGAGCAGGAATGGCCAAAGCTGGCTCTGTGAGGCGGAG 1608
Db 2247 GTTAACCGGAGGCTCAGCTGGAGCAAGAAATGGCCAAAGCTGACTCTGTGAAGCTGAG 2306
QY 1609 CAGGAGCAGATGAGGAAGATCCTCTACCAAGAGAGTCTAACTACAACCGCTCAAGAGG 1668
Db 2307 CAGGAGCAGATGCGGAAGATCCTCTACCAAGAGAGTCTAATTAACAGAGTTAAAGAGG 2366
QY 1669 GCGAAGATGGACAAGTCCATCTTTGTGAAATCAAGACTCTTAGGCATCGGTGCTTTGGG 1728
Db 2367 GCGAAGATGGACAAGTCTATCTTTGTCAAGATCAAAACCCCTGGGATCGGTGCTTTGA 2426
QY 1729 GAAGTGTCCCTCGCTGTAAAGCTGAGCAGCTCAGCTCTGTAGCCCATGAAGACTCTCAGG 1788
Db 2427 GAAGTGTCCCTCGCTGTAAAGCTGAGCAGCTCAGCTCAGCTCTGTAGCCCATGAAGACTCTCAGG 2486
QY 1789 AAGAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGCATCTG 1848
Db 2487 AAAAGGATGCTCTGAACCGGAATCAGTGGCCCACTGAGGCGGAGAGGAGCATCTG 2546
QY 1849 GCTAAGCAGACAATGAGTGGGTGGTCAAACTCTACTCTTCTTCCAGGACAAAGCAGC 1908
Db 2547 GCCAGGCGCAGACAATGAGTGGGTGGTCAAACTCTACTCTTCTTCCAGGACAAAGCAGC 2606
QY 1909 CTGTACTTTGTGATGGACTACATACAGGCGGGGATATGATGAGCCCTGCTGATCAGGATG 1968
Db 2607 CTGTACTTTGTGATGGACTACATCCCTGGTGGGAGATGATGAGCCCTGCTGATCCCGATG 2666
QY 1969 GAGTCTTCCCTGAGCACCTGGCCGCTTCTACATTTGCAGAGTTGACCTTGGCCATTTGAA 2028
Db 2667 GAGTCTTCCCTGAGCACCTGGCCGCTTCTACATCCAGAGCTGACTTTGGCCATTTGAG 2726
QY 2029 AGTGTCCACAAGTGGCTTTATCCACCAGGAGATCAAGCTTCAACATCTCTATCTGAC 2088
Db 2727 AGTGTCCACAAGTGGCTTTATCCACCAGGAGCTGATCCACCGAGACATCAAGCTTGTATAGAT 2786
QY 2089 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGGATTTCAGTGTGACTCAC 2148
Db 2787 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGGATTTCAGTGTGACTCAC 2846
QY 2149 AATTCAGTACTACCAAGAGGAGACCATGAGCAGAGCAGCATGGAGCCCGGTGAC 2208
Db 2847 AATTCAGTACTACCAAGAGGAGACCATGAGCAGAGCAGCATGGAGCCCGGTGAC 2906
QY 2209 CTCTGGCAGCATCTTCCAACTGCTGCTGGAGACAGGTTAAAGACCTTGGAGCAGAGG 2268
Db 2907 CTCTGGGATGATGTCTTAACTGCTGGTGGGAGCAGGCTGAAGACCTTAGAGCAGAGG 2966
QY 2269 GCGCAGAGCAGCAGCAGAGGTGCTTGGCAGATTTCTTGTGGGACACCAATTTACATC 2328
Db 2967 GCGCGAAGCAGCAGCAGAGGTGCTTGGCAGATTTCTTGTGGGAGCTCCAACTACATC 3026
QY 2329 GCTCCCGAGGTGCTTCTCCGCAAGGTTACAGCAGCTCTGTGACTGTGGAGCGTCTGCT 2388
Db 3027 GCACCCGAGGTGCTTCTCCGCAAGGTTACACTCACTGTGACTGTGTGGAGTGTGGA 3086
QY 2389 GTGATTTCTTTTGTGAGATGCTGTTGGGAGCGCGCTTCTTGGCCCGCCACCCACAGAG 2448
Db 3087 GTGATTTCTTTTGTGAGATGCTGTTGGGAGCGCGCTTCTTGGCCCGCCACCCACAGAG 3146
QY 2449 AGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAATATCCCTACGAGGTGAGGCTC 2508
Db 3147 ACCCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAATATCCAGCCAGGTGAAGCTG 3206
QY 2509 AGCGCTGAGGCGCGGAGCCTCATCAGAGCTGTGCTGGCGGCTGACTGCGCGCTGGGCG 2568
Db 3207 AGCGCTGAGGCGCGGAGCCTCATCAGAGCTGTGCTGGCGGCTGACTGCGCGCTGGGCG 3266
QY 2569 AGGATGGGCGGAGTGAACCTCAAGGACACCGCTTCTTCAACACCATCGACTTTTCCGCT 2628
Db 3267 CGGAATGGGCGGAGTGAACCTCAAGGACACCGCTTCTTCAAGGCGGCTTCTTCAAGGCGGCTTCTTCCAGT 3326

```

QY 2629 GATATCCGAAAGCAGGCTGCACCTACGCTCCACCATCAGCCACCCATGACACCTCC 2688
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3327 GATATCCGAAAGCAGCAGCCCTAGCTTCCACCATCAGCCACCCATGACACCTCG 3386
QY 2689 AATTTTGACCGGTGATGAAGAACCCCTGGCAGGAGCCAGCGAGAGCGCCAAAG 2748
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3387 AATTTTGACCGGTGATGAAGAACCCCTGGCAGGATGCCAGGAGGATGACACCAAG 3446
QY 2749 GCCTGGGACACGCTGGCTCCGCCAGCAGCAAGCATCCAGAGCAGCCCTTCTATGATT 2808
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3447 GCCTGGGACACACTCACCTGCCCAATACAAAGCATCTGACACGATTTTACGAATTC 3506
QY 2809 ACCTTCCGAGGTTCTTCGATGACAAACGCTATCCCTCCCGTGGCCGAGCCCTCAG-- 2866
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3507 ACCTTCCGAGGTTCTTCGATGACAAACGCTATCCCTCCCGTGGCCGAGCCTTCAGGA 3566
QY 2867 -----ACCCGCGAGAGTGCAGACCCAGGGGATGCGGATGCGAAGGTGCGGCC 2916
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3567 GCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATGCTGGTGCAGACT 3626
QY 2917 GAGGCTGCCAGCCGCTGACGTGTA 2942
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3627 GAAGCTGCCAGCCCTGTGTACGTGTA 3652

RESULT 4
LOCUS      E38227             5486 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION Human tumor regulatory gene.
ACCESSION  E38227
VERSION    E38227.1 GI:18626936
KEYWORDS   JP 200210086-A/2.
SOURCE     Homo sapiens.
ORGANISM   Homo sapiens.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  1 (bases 1 to 5486)
AUTHORS   Koga,J., Kono,K. and N.Z.F.
TITLE     Human tumor regulatory gene
JOURNAL   Patent: JP 200210086-A 2 02-AUG-2000;
            NIHON CHEMICAL RESEARCH K K
COMMENT    OS Homo sapiens (human)
            PN JP 200210086-A/2
            PD 02-AUG-2000
            PF 25-JAN-1999 JP 1999016223
            PR
            PI JUNICHI KOGA,KEIKO KONO,ZOROTARYOFU FIODORU N PC
            C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68, PC
            G01N33/50,
            PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09, PC
            C12R1:91),C12N15/00.
            CC A61K37/36,(C12N15/00,C12R1:91)
            FH Key Location/Qualifiers
            FT source 1..5486
            FT Location/Qualifiers
            FT 1..5486
            FT /organism="Homo sapiens"
            FT /db_xref="taxon:9606"

BASE COUNT 1380 a 1451 c 1388 g 1267 t
ORIGIN

Query Match 55.4%; Score 1748; DB 6; Length 5486;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 545; Indels 153; Gaps 16;

QY 1 ATGAGAGCACCACCCGAAAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 60
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
576 ATGAGAGCACCACCCGAAAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 635
QY 61 CTCCTGCCTTTTGGCAACAGTACGACCTTCGGCAGCTGACAGGTGAACCGGAGATG 120
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 636 TTGTTGCCCTTTTGTAAATGAATCGGCACTCTGCAGCTGCAGAGTGAACCGCAATG 695
QY 121 CTTTCAGGAGTTTGGTGAATCGGCACTGTACACAGGAGATGGCTGCAGAGCGCTCAGCG 180
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
696 CTGCAGGAAGTGGTGAACGCAAGATGCGACAGGAGATGGCTGGCGGAGCTCTCAAGCAG 755
QY 181 ACGGGCAGTAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAAGATGGCTTACCTGGAC 240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
756 ACTGCGCAGCAGGAGATCGAGGCCGCTGGAGTACATCAGCAAGATGGCTTACCTGGAC 815
QY 241 CCGAGAAATGAGCAGATGTCGAGTCAACAGCACTCCCGAGAAAGGCGCTGGCG 300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
816 CCGAGAAATGAGCAGATGTCGCGGTCAATTAAGCAGACCTCCCGAGAAAGGCGCTCATG 875
QY 301 TCCACCCCGGTACATCGCGGCCAGTTTCGAGGCGACAGGAGACACTCCCATCTTAC 360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
876 CCAACCCCGAGTACGCGGAGGCCAGCTTCGAAGGAACCGCGGATTCGTTTCGCTCTTAC 935
QY 361 CACCACTGGGTGGTCAAACTACGAG-----GGCCCGCCCGCA 399
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
936 CACCACTGAGCGGTACCCCTACGAGGCCCAAGCTTCGGCGCTCAGCGGCCCGCAGGCG 995
QY 400 CTGGAGGAGATGCCGGGCAATATTATGACTTTCTTCCCTGGAGCCGAGCGCCGAC 459
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
996 CTGGAGGAGATGCCGGGCGCTACGTGGACTACCTTTTCCCGGAGTCCGCCCGCCACGGG 1055
QY 460 CACGCTGCCAGGCTCACCAAGGATCTCCCAAGGATACAGCAGCAGTACAGCAAGT 519
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1056 CCGCGCCACGAG--CACCAGCACCACCAAGGCTAGGGTGCAGGCTAGAGGACGCA 1112
QY 520 GCG-----CACTTTCCGGGCACACACTATGTGCTGGTCACTACT-----ATCG 564
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1113 GGGGCACACTTCCCGTCGAGGGCGGCACTACGCGGCGCCGCACTGCTGTGCTGGG 1172
QY 565 GAGCAGTCTGGGTATGGGTGCGAGCGAGTCTCTCTTCCAGAAAGACGACCAAGAT 624
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1173 GAACCCCTGGGTACGGAGTGCAGGCGAGCCCTCTCTTCCAGAGAAAGACGCGCGGAG 1232
QY 625 GCC-----TATTCAGCATGCGC---AAGGCCAGGGTGGGCCCTCC---CGCAGCCTC 672
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1233 ACCGGGGTTACGCGAGCTGCCCAAGAGGCCAGGAGGACCGCGAGCGCCGCGCTC 1292
QY 673 ACCTTTCTGCCCATGCTGGGCTGTACACTGCTCGCACCAAGCGCGGCTACCCCA 732
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1293 GCTTTCCACACCCCTGCGCGCGGCTCTACGTGCGCGACCCACACCAAGCAGCGGT 1352
QY 733 CTTGGGGCCACCCCATATGATGTGTGGGCACCCCGGGTTCACAGTTTACTGGCGAAAGC 792
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1353 CCGGTGGGCCACCACTGCTGTGTGGGCTCCCGCAGCAGGTGTTCGCCAGCAGCAGC 1412
QY 793 TCTGCACAGGCTGCTGTCACCCGTCAGGAACAGCCTCAATGCTGACTGTACAGAGCTG 852
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1413 CCGCGCAGAGCTGCTACTCTCGGGAACAGCTCAAGCTGAGCTGTATGAATG 1472
QY 853 GG-----CTCCAGGTGCGCTTGGTGTGCTGAGTCCACTGGCACCGCGGACTTCGCTG 906
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1473 AGCAGCACTCGTCCAGCAGTGGCGGTGCCACCTTGGCCCGCGGAGTCCCTCGAG 1532
QY 907 AAGCAGGCTCTAGAAGCCTCGCGG---CGCATGTGGCTTTTCGGGCTGGC-----954
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1533 AAGCCGCGCTGGAGGCGCGCGCGCGCTGCGAGTGGGCTTCGCGGCTGACTGCGCAGTG 1592
QY 955 CCGCAGGACCACTCTTCAACAAACCAAC-----CT 990
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1593 CCGCAGGACCACTCTTCAACAGCCACCGCGCGCGCGCTGCGGCTGGCAAGGCC 1652
QY 991 GAGCCCTCACTGCCCGCCCAACACAGGTTCACCGCGTGCAGCGCGCACATCTTCCAC 1050
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1653 GAGCCCTCCCTGCCCGCCCAACACCGTGCAGGCTGTACGCGCGCGCACATCTTGCAC 1712
QY 1051 CTTGTGAAGAGGTGCTGTGCGGCCGAGCCCCAGACAGCGGTGGGGCCCTTCGAC 1110
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1713 CCGGTGAAGAGGTGCTGTGAGGCGGAGCGCGGCTGCTGGGCGCTTCGAC 1772

```


Qy 1111 CCCGCTGGGTGGCTGGC-----CCACAGCACCTGCCACTGAG 1149
 Db 1773 CCCGCTGGGTGGCTGGC-----CCACAGCACCTGCCACTGAG 1832
 Qy 1150 AGCTGGAGACGAAGG-----GGCAGCGGAGGCCACACCCCGTG 1191
 Db 1833 GGCCTGGACGCAAGGAGGAGCATGCCCTGGCTGGCGGCGGAGGCCCTTCCCGTG 1892
 Qy 1192 GATGTGGACTATGGCGCTCCGAGCGCAGGTGCCACCGCTCCGTATCCAAAGCACTTG 1251
 Db 1893 GACGTGGAGTACGAGGCCACAGACCGGAGGTGCCCGCTCCCGCTACCCGAAAGCACCTG 1952
 Qy 1252 CTGCTGCCAGTAAAGTCTGAGCAGTACAGCGTGGACCTGAGAGCCTGTGCACCACTGTG 1311
 Db 1953 CTGCTGCCAGCAAGTCGAGCAGTAC-----GACCTGGACAGCCTGTGCGCAGGCATG 2006
 Qy 1312 CAGCAGAGTCTGCAGGGGG---CACTGATCTAGAGGGAGTACAGAGGCCACAAAGGT 1368
 Db 2007 GAGCAGAGCCTCCGTGGCGGCCCAACAGAGCCCGGAGCGCGCAGAGCGGCAAAAGC 2066
 Qy 1369 GCGAAGGAGACAAGCTGGCGAGAGACAAAAGCAGATTACAGACTCCCGGTGCCCTGTC 1428
 Db 2067 GCCAAGGGGACAAAGCGGAAAGATATAAAGCAGATTACAGACTCTCCCGTTCGCGTTC 2126
 Qy 1429 CGCAAGATAGCAGAGATGAAGAGAAGAGAGAGTCTCGCATCAAGAGTACTCCCTTAT 1488
 Db 2127 CGCAAAAACACAGAGAGAGAGAGAGAGTCAAGCATCAAGAGTACTCGCCATAC 2186
 Qy 1489 GCCTTCAATCTTCATGAGACACACAGTGGAGATGTATCAAAACCTACACAGAG 1548
 Db 2187 GCCTTTAAGTCTTCATGAGCAGCAGCTGGAGATGTATCAAAACCTACACAGAG 2246
 Qy 1549 GTACGCGGAGGCTACAGCTGAGCAGAGAAATGCCAAAGCTGGCTGTGAGGCGCAG 1608
 Db 2247 GTTAACCGAGGCTGACAGTGGAGCAAGAAATGCCAAAGCTGGACTCTGTGAAGCTGAG 2306
 Qy 1609 CAGGAGCAGATGAGGAAGATCCTTACCAGAAAGAGTCTTAACACCGGCTGAAGAGG 1668
 Db 2307 CAGGAGCAGATGCGGAAGATCCTTACCAGAAAGAGTCTTAACACAGGTTAAAGAGG 2366
 Qy 1669 GCCAAGATGACAGTCCATGTTTGTGAATCAAGACTCTAGCAGCTGGTCCCTTTGGG 1728
 Db 2367 GCCAAGATGACAGTCCATGTTTGTGAATCAAAACCTGGGATCGGTCTTTTGA 2426
 Qy 1729 GAAGTGTCCCTGCTGTGAAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGG 1788
 Db 2427 GAAGTGTCCCTGCTGTGAAGCTGGACACTCACGCTCTGTACGCCATGAAGACCTTAAG 2486
 Qy 1789 AAGAAGATGTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 1848
 Db 2487 AAAAAGATGTCTGAACCGGAATCAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 2546
 Qy 1849 GCTGAACGACAGATGAGTGGTGGTCAAACTCTACTCTCTTCCAGGACAGGACAGC 1908
 Db 2547 GCCAGGACAGATGAGTGGTGGTCAAACTCTACTCTCTTCCAGGACAGGACAGC 2606
 Qy 1909 CTGTACTTTGTGATGAGTACATACAGGCGGGATATGATGAGCCTGTGATCAGGATG 1968
 Db 2607 CTGTACTTTGTGATGAGTACATACAGGCGGGATATGATGAGCCTGTGATCAGGATG 2666
 Qy 1969 GAGTCTTCCCTGAGCAGCTGGCCCGTCTTACATTTGACAGTTGACCCCTGGCCATTGAA 2028
 Db 2667 GAGTCTTCCCTGAGCAGCTGGCCCGTCTTACATTTGACAGTTGACCCCTGGCCATTGAG 2726
 Qy 2029 AGTCTCACAGATGGGCTTTATCCACCGGACATCAAGCCTGACACATCTACTCATGAC 2088
 Db 2727 AGTCTCACAGATGGGCTTTATCCACCGGACATCAAGCCTGATATCTTTGATAGAT 2786
 Qy 2089 CTGATGGTCTATTAAGCTGACAGATTTTGGCCTCTGCACTGGATTCAGGTGAGCTAC 2148
 Db 2787 CTGATGGTCTATTAAGCTGACAGATTTTGGCCTCTGCACTGGGTTTCAGGTGAGCTAC 2846

Qy 2149 AATTCCAAGTACTACCAAGAAAGGAAACACATGAGACAGGACATGAGACCCCGGTGAC 2208
 Db 2847 AATTCCAATATACCAAGAAAGGAGCCATGTACACAGGACAGCATGAGACCCCGGAC 2906
 Qy 2209 CTCTGGGACGATGTTTCCAACTGCTGCTGGACAGGTTAAAGACCTCGAGCAGAGG 2268
 Db 2907 CTCTGGGATGATGTTTAACTGCTGCTGGGACAGGCTGAAGACCTTAGACAGAGG 2966
 Qy 2269 GCGCAGAAAGCAGCAGCAGAGGTGCTGSCACATCTCTTGTGCGGACACCAAAATACATC 2328
 Db 2967 GCGCGAAGCAGCAGCAGAGGTGCTGSCACATCTCACTGGTGGGACTCCAAACTACATC 3026
 Qy 2329 GCTCCGAGGTGCTTCTCCGAAAGGTTACACGAGCTCTGTGACTGTGTGAGGAGGTCGGT 2388
 Db 3027 GCACCGAGGTGCTTCTCCGCAAAAGGTTACACTCAACTCTGTGACTGTGTGAGTGTGGA 3086
 Qy 2389 GTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTTCTGGCCCCCACCACACAGAG 2448
 Db 3087 GTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTTCTGGCACCTTACTTCCACAGAA 3146
 Qy 2449 ACGCAGCTGAAGGTGATCACTGGGAGAGCAGCTGCATATCCCTACGAGGTGAGGCTC 2508
 Db 3147 ACCCAGCTGAAGGTGATCACTGGGAGAGCAGCTGCATATCCAGCCAGGTTGAAGCTG 3206
 Qy 2509 AGCCTGAGGCGCCGAGACCTCATCAGAAAGCTGTCTGCGGCGGTGACTGCCGCTGGGC 2568
 Db 3207 AGCCTGAGGCGCCGAGACCTCATCAGAAAGCTGTCTGCGGAGACCTGAGGCTGGG 3266
 Qy 2569 AGGATGGGCGAGATGACCTCAAGGCACACCCGTTTCTTCAACACCATCGACTTTCCGCT 2628
 Db 3267 CGGAATGGGCGGATGACCTGAAGGCCACCCCTTCTTCAAGCGCATTTGACTTCTCCAGT 3326
 Qy 2629 GACATCCGAAAGCAGGCTGCACCTACGCTCCCAACCATCAGCCACCCATCGACACCTCC 2688
 Db 3327 GACATCCGAAAGCAGGACGACCCCTACGCTCCCAACCATCAGCCACCCATCGACACCTCG 3386
 Qy 2689 AATTTGACCCGCTGGATGAAGAAAGCCCTTGGCACAGGCGCAGCGAGAGAGGCCAAG 2748
 Db 3387 AATTTGACCCGCTAGATGAAGAAAGCCCTTGGAAACGATGCCAGGAAGGTAGCACCAAG 3446
 Qy 2749 GCCTGGACACAGCTGGCTCCCGCAGCAGCAGCATCCAGAGCAGCGCTTCTATGAGTTC 2808
 Db 3447 GCCTGGACACACTCACTCGCCCAATAAAGCATCTCTGAGCAGCATTTTACGAATTC 3506
 Qy 2809 ACCTTCCGAGGTCTTCTGATGACAGCGGTATCCCTTCCGCTGCCGAGGAGCCCTCAG-- 2866
 Db 3507 ACCTTCCGAGGTCTTCTGATGACAGGTATCCCTTCCGCTGCCGAGGAGCCCTCAGGA 3566
 Qy 2867 -----AGCCCGCAGAGAGTGCAGACCCAGGGATCGGACTTTGGAAGGTGCGGCC 2916
 Db 3567 GCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3626
 Qy 2917 GAGGGCTGCCAGCGGTGTACGTGTA 2942
 Db 3627 GAGGGCTGCCAGCGGTGTACGTGTA 3652

RESULT 5
 AF207547
 LOCUS AF207547
 DEFINITION Homo sapiens serine/threonine kinase KPM mRNA, linear PRI 01-JUL-2000
 ACCESSION AF207547
 VERSION AF207547.1 GI:8886763
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4098)
 AUTHORS Hori, T., Takaori-Kondo, A., Kamikubo, Y. and Uchiyama, T.
 TITLE Molecular cloning of a novel human protein kinase, kpm, that is homologous to warts/lats, a Drosophila tumor suppressor
 JOURNAL Oncogene 19 (27), 3101-3109 (2000)

```

MEDLINE 20332247
PUBMED 10871863
REFERENCE 2 (bases 1 to 4098)
AUTHORS Hori,T.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-1999) Hematology and Oncology, Kyoto University,
54 Kawaracho, Shogoin, Sakyo-ku, Kyoto, Kyoto Prefecture 606-8507,
Japan

FEATURES
source
Location/Qualifiers
1..4098
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="myeloid cell line"
375..3641
/note="similar to Drosophila warts/lats"
/codon_start=1
/product="serine/threonine kinase KPM"
/protein_id="AAF80561.1"
/db_xref="GI:8886764"
/translation="MRPKTPATYTSNSRRLQEIIRGLKQPSKSSVQGLPAGPNSD
TSLDAKVLGSKDARQQOMRATKFGPYOKALREIRYSLLPANESQVTSAAAEVNRQ
MLQELVAGCDEMGARALQKOTGRSIEAALYISKWYLDPRNEQIVRVTQKSPGK
GLMPTVTRPSPEGTGDSFASYHQLSGTPVEGSPFGADGPTALEEMPRVVDYLFPG
VGHPGCHQHQPCKGYSAGVEAGAHFPLQAGYGRPHLLVPEPLGYGVORSPFSD
SKTPEPTGASLTPTKGGGPPGAGLAPPPAAGLYVPHPHKQAGPAAHQHLVIGSR
SQVFASDPQSLLTPSRNSLDVLYELSSVQVQPAATLARRDSLOKPLEAPRA
HVAFRPDCVPVSRNFSNHPGPPGKAPSLPAPNTVTAHAHLHPKSVRL
RPEQTAVGSHPAWPAAPAPAPAPAAEAGLDAAKEHALALGGAGAPLDVEYGG
PDRRCPPVPKHLRLSKSEYDLSLACAGMEOSLRAGNPEPSGDKSRKSGDKG
GDKKQIOTSPVVRKNSRDEEKRESRIKSYSPYAFKPFMEQHVENVIKTYQKVRNR
LQLEKNAKGLCAEBOEMKILYQKESNRYNRUKRAKMDKSMFEVKIKTLGIFAGEV
CLACKVTHALYAMKTRKLDVNRQVHVKAERDILAEADNEWVKLYSFQDKDS
LYFVMDYIPGDMMSLLIRMEVPEFLAREYIAELTIAESVHKMGTHIRKPDNLT
IDLGHKILDFGLCTGRWTHNSKYQKSHVRQDSMEPSDLMDDVSNCRGDLRT
LEQRARQHORLHSLVGPVNYTAPEVLLKGYTOLCDMWSVGLIFELMLVGOPPEL
APTPTQLKVINVENTLHPIPAQVLLSPEARLDLTKLCCSADHRLGRNADDLKHPF
FSALDFSDIRKQAPAPVPTISHPMDTSNDFVDESPFNDASBEGSNAMDTLSPNN
KHPHAYEYETFRFFDDNGYPPFRPKPSGAESAQAESDLSLDVDTQEGCOPVTV
BASE COUNT 1006 a 1218 c 1130 g 744 t
ORIGIN

Query Match 55.48; Score 1746.4; DB 9; Length 4098;
Best Local Similarity 77.38; Pred. No. 0;
Matches 2387; Conservative 0; Mismatches 546; Indels 153; Gaps 16;

Qy 1 ATGAGAGCACCCGAGCTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATTC 60
Db 564 ATGAGAGCACCCGAGCTTGGACCTTATCAGAAGCCTTGAGGAAATCAGATATTC 623

Qy 61 CTCCTGCTTTTGGCAACGAGTCAGGCACCTTCGGCAGCTGCAGAGGTGAACCGGAGATG 120
Db 624 TTGTTGCCCTTTTGTGTAATGAATCGGGCACCTCTGCAGCTGCAGAGTCAACCGGCAATG 683

Qy 121 CTTAGGAGTTGGTAATGCGGCATGTACACAGAGATGGCTGGCAGAGCGCTACGAGC 180
Db 684 CTGAGGAACCTGGTGAACGCGAGGATGCGACGAGGAGATGGCTGGCCGAGCTCTCAAGCAG 743

Qy 181 ACGGCGAGTAGGAGTATCGAAGCTGCCPTTGGAGTACATCAGTAAGATGGCTACTGTGAC 240
Db 744 ACTGGCAGCAGAGCATCGAGCGCCCTCTGGAGTACATCAGCAAGATGGGCTACTGTGAC 803

Qy 241 CCCAGGAATGACGAGATGTGCGAGTCAATCAAGCAGACCTCCCGAGGAAGGGCTGCGC 300
Db 804 CCGAGGAATGACGAGATGTGCGGCTCATTAAGCAGACCTCCCGAGGAAGGGCTCATG 863

Qy 301 TCCACCGCGTACTCGCGCCCGCATTTTCGAGGCGACAGGGGAGCACTCCCATCTTAC 360
Db 864 CCAACCCCGAGTGAACGCGGAGCCAGCTTCGAAGGAACCGCGGATTCGTTTTCGCTCTAC 923

Qy 361 CACCAGCTGGGTGGTGAACACTACGAG-----GGCCCCCGCCCA 399
Db 924 CACCAGCTGAGCGGTGACCCCTACGAGGGGCCCAAGCTTCGCGGTGAGCGGCCCAACGCGC 983

```

```

Qy 400 CTGGAGGAGATGCCGCGCAATATTAGACTTTCTTCCCTGGAGCGGAGCCGCGCACC 459
Db 984 CTGGAGGAGATGCCGCGCGCGCTACGTGGACTACCTTTTCCCGGAGTGGCGCCCGCAGGG 1043
Qy 460 CACGGTGCCCGGCTCACACAGCATCTCTCCAAAGGTTACAGCACAGCATGAGACCAAGT 519
Db 1044 CCGCGCCACAG---CACCAGCACCCACCAAGGGCTAGGGTCCAGCGCTAGAGCAGCA 1100
Qy 520 GCG-----CACTTCCCGGCGACACACTATGCTGGTGTCTACTACT-----ATCG 564
Db 1101 GGGGACACATTCCTCCGCTGAGGGCGCGCACTACGGGCGCGCGCACTGCTGTGTGCTGGG 1160
Qy 565 GAGCAGTCTGGGTATGGGTGCGAGCGAGTCTCTCTCTCCAGAAAGACGACGACCAAGAT 624
Db 1161 GAACCCCTGGGTACGAGGTGCGAGCGACGCCCTCTCTCCAGAGCAAGACGCGCGGAG 1220
Qy 625 GCC-----TATTCACAGATGGCC---AAGGCCCCAGGTGGCCCTTCC---CGCCAGCCCTC 672
Db 1221 ACCGGGGGTTAGCGCAGCTGCCCCACGAAGGCGCAGGAGGACCGCCAGGCGCGCGCTC 1280
Qy 673 ACCTTTCTGCTGCTGCTGGGCTGTACACTGCTCGCACCAAGCGCGCGGCTTACCCCA 732
Db 1281 GCTTTTCCACCCCTGCGCGCGGGGCTCTACGTGCGCGCACCCACACCAAGCAGCGCGGT 1340
Qy 733 CCTGGGGCCCCACCATTTACATGTTGGGCACCCCGGGTCCCACGTTTACTTGGCGAAAGC 792
Db 1341 CCGCGCGCCCCACCATGCTGCTGCTGGCTCCCGCAGCGAGGTGTTGCCAGCAGCACAGC 1400
Qy 793 TCTGCACAGGTGCTGCTGGCACCTGCCAGGACACGCTCAATGCTGCTGTGTGACGAGCTG 852
Db 1401 CCGCGCGAGAGCTGCTCTACTCTCCCTCGGGAACAGCTCAAGCTGAGCTGTGTAATG 1460
Qy 853 GG-----CTCCACGCTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
Db 1461 AGCAGCAGCTTCCGTCACGAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1520
Qy 907 AAGCAGGCTCTAGAAAGCTTCGCGGC---CGCATGTGGTGTGTTTTCGGGCTGGC-----954
Db 1521 AAGCGGGGCTTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1580
Qy 955 CCGCAGCAGCACTCTCTTCAACAAACCCACAC-----CT 990
Db 1581 CCGCAGCAGCACTCTCTTCAACAGCCAGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1640
Qy 991 GAGCCTCAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1050
Db 1641 GAGCCTCCTCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1700
Qy 1051 CCGTGAAGAGCGTGCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1110
Db 1701 CCGGTGAAGAGCGTGCCTGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1760
Qy 1111 CCGCGCTGGGTGGCTGCGC-----CCACAGCAGCTGCGCAGT 1149
Db 1761 CCGCGCTGGGTGGCTGCGC-----CCACAGCAGCTGCGCAGT 1820
Qy 1150 AGCCTGGAGAGCAGGA-----GGCAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191
Db 1821 GCTTGGAGCGCAGGAGGAGGAGTCCCTGCGCTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 1880
Qy 1192 GATGTGGACTATGGCGGCTCCGAGCGAGGTGCCACCGCTTCCCTATTCCTCAAGACATG 1251
Db 1881 GACGTGGAGTACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1940
Qy 1252 CTGCTGCCCACTAGTCTGAGCAGTACAGCTGAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAG 1311
Db 1941 CTGCTGCCCACTAGTCTGAGCAGTACAGCTGAGCAGCTGAGCAGCTGAGCAGCTGAGCAGCT 1994
Qy 1312 CAGCAGAGTCTGCGAGGGGG---CACTCATCTAGACGGGAGTACAGCAGCAGCAGCAGCAGCAG 1368
Db 1995 GAGCAGAGCCTTCCCTGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2054

```

QY 1369 GCGAAGGAGACAAAGCTGCGCAGAGACAAAAGCAGATTTCAGACCTCCCGGCTGCCTGTC 1428
 Db 2055 GCCAAGGGGACAAAGCGGAAAGGATAAAAGCAGATTTCAGACCTCTCCCGTTCCCGTC 2114
 QY 1429 CGCAAGATAGCAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTAT 1488
 Db 2115 CGCAAAAGCAGAGAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTATC 2174
 QY 1489 GCCTTCAAAATCTTCATGGAGCAGACGCTGGAGAATGTCAATCAAAACCTACCAAGCAGAG 1548
 Db 2175 GCCTTAAAGTTCTTCATGGAGCAGACGCTGGAGAATGTCAATCAAAACCTACCAAGCAGAG 2234
 QY 1549 GTCAGCGGAGGCTACAGCTGGAGCAGAGAAATGCCAAAGCTGGGCTCTGTGAGCGCGAG 1608
 Db 2235 GTTAAACCGGAGGCTGCAAGCTGGAGCAGAGAAATGCCAAAGCTGGGCTCTGTGAGCGCGAG 2294
 QY 1609 CAGCAGCAGATGAGAGAGATCTCTACCAAGAGGCTTAATCAAAACCGGCTGAAGAGG 1668
 Db 2295 CAGCAGCAGATGCGGAAGATCTCTACCAAGAGGCTTAATCAAAACCGGCTGAAGAGG 2354
 QY 1669 GCCAAGATGGAACAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 1728
 Db 2355 GCCAAGATGGAACAGTCTATGTTGTCAAGATCAAAACCGGCTGGGATCGGTGCTTTGGA 2414
 QY 1729 GAAGTGTGCTGCTGTAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGG 1788
 Db 2415 GAAGTGTGCTGCTGTAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGG 2474
 QY 1789 AAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 1848
 Db 2475 AAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 2534
 QY 1849 GCTGAAGCAGACATAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGGACAGC 1908
 Db 2535 GCCAGGAGCAGACATAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGGACAGC 2594
 QY 1909 CTGTACTTTGTGTGAGTACATACAGCGGGGATATGATGAGCCTGCTGATCAGGATG 1968
 Db 2595 CTGTACTTTGTGTGAGTACATACAGCGGGGATATGATGAGCCTGCTGATCAGGATG 2654
 QY 1969 GAGGTCTTCCCTGAGCAGCTCGCCGCTTCTACATTTGAGAGATTTGACCTGGCCATTTGAA 2028
 Db 2655 GAGGTCTTCCCTGAGCAGCTCGCCGCTTCTACATTTGAGAGATTTGACCTGGCCATTTGAG 2714
 QY 2029 AGTGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGTCAACACTCTATCTCAC 2088
 Db 2715 AGTGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGTCAACACTCTATCTCAC 2774
 QY 2089 CTGGATGCTCATATTAAAGCTCAGAGATTTTGGCCTCTGCACTGGATTGAGTGGACTCAC 2148
 Db 2775 CTGGATGCTCATATTAAAGCTCAGAGATTTTGGCCTCTGCACTGGATTGAGTGGACTCAC 2834
 QY 2149 AATTCAGATCTACCAAGAGGAAACACATGAGACAGGACAGCATGGAGCCGGTGAC 2208
 Db 2835 AATTCAGATCTACCAAGAGGAAACACATGAGACAGGACAGCATGGAGCCGGTGAC 2894
 QY 2209 CTCTGGAGGATGTTTCAACTGTCTGCTGGAGACAGGTTAAAGACCTCTGGAGCAGAGG 2268
 Db 2895 CTCTGGAGGATGTTTCAACTGTCTGCTGGAGACAGGTTAAAGACCTCTGGAGCAGAGG 2954
 QY 2269 GCGCAGAGACAGCAGAGGCTGCTGGACATTTCTTTGCGGAGACACCAAAATACATC 2328
 Db 2955 GCGCAGAGACAGCAGAGGCTGCTGGACATTTCTTTGCGGAGACACCAAAATACATC 3014
 QY 2329 GCTCCGGAGGCTCTTCCGCAAGAGGTACAGCAGCTGTGACTGGTGGAGCGCTCGGT 2388
 Db 3015 GCAACCGAGGCTCTTCCGCAAGAGGTACACTCAACTCTGTGACTGGTGGAGCTTGA 3074
 QY 2389 GTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTCTTTGGCCCCCAGCCACAGAG 2448
 Db 3075 GTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTCTTTGGCCCCCAGCCACAGAA 3134
 QY 2449 ACGCAGCTGAAGGTGATCAACTGGGAGCAGCGCTGCATATCCCTACGAGGTGAGGCTC 2508

Db 3135 ACCAGCTGAAGTGTATCAACTGGGAGACACGCTCCACATTCAGCCCGAGGTGAAGCTG 3194
 QY 2509 AGCGTGAAGCGGCGAGACCTTCATCAGGAAGCTGTGCTGCGCGCTGACTGCGCCTGGGC 2568
 Db 3195 AGCCCTGAGCGAGGAGACCTTCATCACCAGCTGTGCTGCTCCGAGACACCGCCTGGGG 3254
 QY 2569 AGGATGGGCGCAGATCAGCTCAAGGCACACCGCTTCTTCAACACCATCAGCTTTTCCCGT 2628
 Db 3255 CGGAATGGGCGCGATGACCTGAAGGCCACCCCTTCTTTCAGCGCCATTCAGTCTCCACT 3314
 QY 2629 GACATCCGAAAGCAGGCTGACCTACCTTACCTCCCAACCATCAGCCACCCATGACACTCC 2688
 Db 3315 GACATCCGAAAGCAGGCGAGCCCTTACCTTCCCAACCATCAGCCACCCATGACACTCCG 3374
 QY 2689 AATTTGACCCCGTGTGATGAAGAAACCCCTTGGACAGGCGGAGGAGCGCAAG 2748
 Db 3375 AATTTGACCCCGTGTGATGAAGAAACCCCTTGGACAGGCGGAGGAGGAGCGCAAG 3434
 QY 2749 GCTGGGACACGCTGCGCTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2808
 Db 3435 GCTGGGACACACTCAGCTCGCCCAATACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3494
 QY 2809 ACCTTCGCGAGGTTCTTCGATGACAGCGCTATCCCTTCGCTGCGGTCGCGAGCCCTCAG 2866
 Db 3495 ACCTTCGCGAGGTTCTTCGATGACAGCGCTATCCCTTCGCTGCGGTCGCGAGCCCTCAG 3554
 QY 2867 -----AGCCCGCAGAGACTGCGAGCCCGGAGGAGTGGAGGCTTGGAGGTCGGGCG 2916
 Db 3555 GCAGAGCTTTCAGAGCTGAGAGCTCAGATTAGAAAGCTCTGATCTGGTGGATCAGACT 3614
 QY 2917 GAGGCTGCGCAGCGGCTGCTACGTGA 2942
 Db 3615 GAAGCTGCGCAGCGTGTGCTAGCTGTA 3640

RESULT 6
 AB028019 3412 bp mRNA linear PRI 09-MAR-2000.
 LOCUS
 DEFINITION Homo sapiens HSLATS2 mRNA for large tumor suppressor 2, partial cds.

ACCESSION AB028019
 VERSION AB028019.1 GI:7212789
 KEYWORDS large tumor suppressor 2; HsLats2.
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)
 Yabuta, N., Fujii, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Nishiguchi, H., Endo, Y., Toji, S., Tanaka, H., Nishimune, Y. and Nojima, H.
 Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts
 Genomics 63 (2), 263-270 (2000)
 20139436
 2 (bases 1 to 3412)
 Nojima, H. and Nishiguchi, H.
 Direct Submission
 Submitted (27-MAY-1999), Hiroshi Nojima, Research Institute for Microbial Diseases, Osaka University, Department of Molecular Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
 (E-mail: nojima@biken.osaka-u.ac.jp, tel: 81-6-6875-3980, Fax: 81-6-6875-5192)

FEATURES
 source Location/Qualifiers
 1..3412
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1..3412
 /gene="HsLATS2"
 <1..3143
 /gene="HsLATS2"
 /codon_start=3

/product="large tumor suppressor 2"
/protein_id="BAA92381.1"
/db_xref="GI:7212790"
/translation=" NSDTSLDAKVLGSKDATSKQOMRATPKFPYQKALREIRSLYLL
PFANESGAAAEVNRQMLQELVNAQDQEMAGRALQKTCGRSIEAALEVIKMGVLL
PNEQTVIRIKYDTSPGKGLMPTVTRRPSFEFGDGSFASYHOLSGPYPEGSPFGADGP
TAALEENRPXYVDLFPFGVGHGQHQPHPKGYGSAVEAAGAHFPLOQAGHYGRPHLL
VCEPGLYVQRPSPFSQKTPPETGYASLTPKGGGPPCAGLAEPPPAAGLYVPHPH
HKAQAPVANHQLVHGSRISOVFADSPQSLITPSSRNLVNLYELSSYSYQOWPATL
ARDSLQKGLAEAPRAHVAFRDCVPVSRKTNFSNHPQRPFGPKAEELSPAPNTVT
AVSAHILHPFKSVSRVLRPEQTAVGSPHAWPAPAPAPAPAPAEGLDAKEEHA
LAUGGAAFLDVEYQGPDRSKALRTRTCCCAKSEQYDLDSLCAGMQSLRAGNGE
PEGGDSRSKAGDKGDKKQDTQTSFVPRVNRNDEEKRESIRKISYSYPAKPFMEQ
HVENVITKTPQGNRRRLQEQEMAKAGLCAEQEOMRKLIYQKESNRLKRAKMDKS
MEFYKLTGFAGFCEYCLACKVDYDTHALYAMKTLRKDLVNRQVHAHVKEATLAEAF
NENWVKLYISFGDKSLSYFMVDYIPGDSMISLLIRMEVFHELAREYIAELTALAEV
HKMFTHRIDIKPDNILLIDGDKLTDFCLGTGFRTKSHSKYQKSGHVQDSMEFSD
LHDDVSNRCGDBELTEQARAKQORSIAHSLVGPNTIAPVELKRGKITQLCDWWS
VGYLLEMGVPPFPFPAETPTETQKLIWNENTLHPIPAQVKLSPEARDLITKLCCSAD
HRUGRGAAOGLKAKHPFSAIDFSSDIRKQAPYVPVTFISHPMDTNSPVDYDEESPNWA
SEGSTRKADMTLSPNNKHPEHAFYEYFRFRFDNDNGYPPRCRCPKSGAEASQAESDLE
SSDLVDQTEGCCQFPVY"
3364..3369
polyA_signal

BASE COUNT
ORIGIN

801 a 1066 c 975 g 570 t

/genes=HSLATS2

Query Match 54.4%; Score 1717.2; DB 9; Length 3412;
Best Local Similarity 77.2%; Pred.No. 0;
Matches 2382; Conservative 0; Mismatches 548; Indels 156; Gaps: 18;

QY 1 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAAGCTCTCAGGGAAATCCGATATTC 60
DB |||||
QY 61 CTCTGCCTTTTGCAACGAGTTCAGGCACCTTCGAGGTGAACCGGCAGAT 120
DB |||||
QY 129 TTGTTCCTTTTGTATGAATCGGCACCCTCTGCAGTGCAGAAAGTGAACCGGCAATG 188
DB |||||
QY 121 CTTTCAGAGTGTGTGAATCGGCATGTGACACGAGATGCTGGCAGAGCGCTCACGCAG 180
DB |||||
QY 189 CTCGAGGACTGTGTACGCGAGATGCGACGAGATGCTGGCGAGCTCTCAAGCAG 248
DB |||||
QY 181 ACGGGAGTAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAAGATGGGTACCTGGAC 240
DB |||||
QY 249 ACTGGCAGCAGGATCGAGGCCGCCCTGGAGTACATCAGCAAGATGGCTACCTGGAC 308
DB |||||
QY 241 CCAGGNAATCAGCAGATTGTGCAGTATCAACGACACTCCCCAGGAAGGCGCTGGCG 300
DB |||||
QY 309 CCAGGNAATCAGCAGATTGTGCGGGTCAATAAGCAGACCTCCCACGGAAGGGGCTCATG 368
DB |||||
QY 301 TCCACCCCGGTGACTCGCGGGCCAGTTTCGAGGGCACAGGGCAAGCACTCCCATCTAC 360
DB |||||
QY 369 CCAACCCAGTGACGGAGGCCAGCTTCGAAGAACCGGCGATTCGTTGGCTCCTAC 428
DB |||||
QY 361 CACCACTGGGTGGTGCACAACTACGAG-----GGCCCCCGCGCA 399
DB |||||
QY 429 CACCACTGACGGTACCCCTACGAGGGGCCAAGCTTGGCGCTGACGGCCCCACGGCG 488
DB |||||
QY 400 CTGGAGGATGCCGGGCAATATTAGACTTTCTTCTGCGAGCGGAGCGCGGAC 459
DB |||||
QY 489 CTGGAGGATGCCGGGCGCGTACGTGGACTACTTTTCCCGGAGTCGGCCCCACGGG 548
DB |||||
QY 460 CAGGTGCCAGGCTCACAGCATCTCCCAAAGGTFACAGCACACGCTAGAGCCAAGT 519
DB |||||
QY 549 CCGGGCACCG---CACACGACACCCACCGAGGGTACGGTGCCAGCTAGAGGACGA 605
DB |||||
QY 520 GCG-----CACTTTCGGGCACACACTATGTGCTGGTCTACT-----ATCG 564
DB |||||
QY 606 GGGGCACACTTCCCGCTGCAGGCGCGECATACGGCGCGGCTGCTGGTGGCTGGG 665
DB |||||
QY 565 GACAGCTTGGGTATGGGTGACGCGAGTTCTCTCTTCAGAACAGACGCCACAGAT 624
DB |||||

Db	666	GAACCCCTGGGCTACGGAGTGCAGCGCAGCCCTCTCTTCCAGAGCAAGACGCGCCGGAG	725
QY	625	GCC-----TATTCCAGCATGCC---AAGGCCCAGGCTGGCCCTCC---CGCCAGCGCTC	672
Db	726	ACCGGGGTTACGCCAGCCTGCCAGGAAGGCCAAGGAGACCGCAGCGCCCGGCTC	785
QY	673	ACCTTTCTCGCCATGCTGGGTGTACACTGCTCGCACCAACAAGCGCGGCTTACCCCA	732
Db	786	GCTTTCCACCCCTCGCCGGGCTCTAGCTGGCGCACCCACACCAAGCAGCGCGGT	845
QY	733	CTTGGGGCCCAACCCATTACATGTTGGGCACCCGGGCTCCACGTTTACTGGCGAAGC	792
Db	846	CCCGTGGCCCAACAGCTGATGCTGGGCTCCGACGCCAGGTGTTTCGCCAGCAGACG	905
QY	793	TCGTGCACAGGCTGTGTGGCACCGTCCAGGAACAGCTCAATGCTGACTTGTACGAGCTG	852
Db	906	CCCCCGCAGAGCTGCTACTCCCTCGCGGAACAGCTCAACGTGGACCTGTATGAATTG	965
QY	853	GG-----CTCCACGGTGCCTTGGTGTGCAGCTCCACTGGCACGCCGCACTCGCTCGAG	906
Db	966	AGCAGCACCTTCGTCAGCAGTGGCGGCTGCCACCTTGGCCCGCCGCGGACTCCCTCGAC	1025
QY	907	AAGCAGGTTAGAACCTCGCGGC---CGCATGTGGCTTTTCGGGCTGGC-----954	954
Db	1026	AAGCCGGGCTGGAGCGCGCGCGCGCAGCTGGGCTTTCCGGGCTGACTGCCCAAGT	1085
QY	955	CCCAGCAGACCAACTCTCTTCAACAACCCACAAC-----CT990	990
Db	1086	CCCAGCAGACCAACTCTTCAACAGCACCAAGCCGCGGCGCGCTCGCGCTGGCAGGCC	1145
QY	991	GAGCCCTCACTGCCCGCCCCAACACGGTCAACCGCGGTGAGCGGCCGCACACATCTTCAC	1050
Db	1146	GAGCCCTCCTCGCCGCCCCCAACACCGTGCAGCGGTGTACGCGCGCGCACATCTTGCAC	1205
QY	1051	CTGTGAAGAGCGTGGTGTCTGCGGCCGAGCCCCAGACAGCCGTGGGGCCCTGCGAC	1110
Db	1206	CCGGTGAAGAGCGTGGTGTGTGAGGCGGAGCGCACAGCGGCTGTGGGGCCCTGCGAC	1265
QY	1111	CCCGCTGGGTGGCTGCGC-----CCACAGCACCTGCACACTGAG1149	1149
Db	1266	CCCGCTGGGTGCCCGGCTGCCCGGCCCGCCCGCCCGCCCGCGGCTGGGAG1325	1325
QY	1150	AGCCTGGAGACGAAGA-----GGCAGCGAGCGCCCCACACCCCGCTG1191	1191
Db	1326	GGCTTGGAGCCCAAGGAGGAGCATCCCTGGCGTGGGCGCGCAGCGGCTTCCCGCTG1385	1385
QY	1192	GATGTGGACTATGGGGCTCCGAGGCGAGTGGCCACCGGCTCCGTATCAAGACACTTG1251	1251
Db	1386	GAGTGGAGTACGGAAGGCCAGACCGGA-GTGCCCGCTCCGCGCTACCCGAAGACACTG1444	1444
QY	1252	CTGCTGCCCAAGTCTTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCACCACTGTG1311	1311
Db	1445	CTGCTGCGCA--AAGTCGAGCAGTAC-----GACCTGGACACGCTGTGCGCAGGCATG1496	1496
QY	1312	CAGCAGAGTCTCCGAGGGG---CACTGATCTAGACGGGAGTGACAAAGCCACAAGGT1368	1368
Db	1497	GAGCAGAGCTCCGTGGGGCCCCAACGAGCCGAGGGGCGCAGACAGAGCCCAAAAGC1556	1556
QY	1369	GCGAAGGAGACAAAGCTGGCAGAGACAAAAGCAGATTACAGCTCCCGGTGCTGTC1428	1428
Db	1557	GCCAAAGGGGACAAAGGGGAAAGGATAAAAGCAGATTACAGACTCTCCCGCTTCCCGTC1616	1616
QY	1429	CGCAAGATPAGCAGAGATTGAAGAGAAGACAGTCTCGCATCAAGAGTACTCCCTTAT1488	1488
Db	1617	CGCAAAACAGCAGACGAAAGAGAGAGAGTACGCAATCAAGAGCTACTCGGCATAC1676	1676
QY	1489	GCCTTCAATTCTTACGTGAGCAACACGCTGGAGAAATGTCATCAAAACCTTACCAGCAGAAG1548	1548
Db	1677	GCCTTTAAGTTCTTATGAGCAGCAGCTGGAGAAATGTCATCAAAACCTTACCAGCAGAG1736	1736
QY	1549	GTCAGCCGAGGCTTACGCTGGAGCAGGAATGCCCCAAAGCTGGGCTCTGTGAGGCCGAG1608	1608
Db	1737	GTTAACCGGAGCTGCAGCTGGAGCAAGAAATGCCAAAGCTGACCTCTGTGAAGCTGAG1796	1796

QY 1609 CAGGACGATGAGGAGATCTCTTACCAAGAGTCTAACTACAACCGCTGAAGAGG 1668
DB 1797 CAGGACGATGCGGAGATCTCTTACCAAGAGTCTAACTACAACCGCTGAAGAGG 1856
QY 1669 GCCAAGATGACAAAGTCCATGTTTGTGAATCAAGACTCTAGGATCGGTCCTTGGG 1728
DB 1857 GCCAAGATGACAAAGTCTTATGTTTGTGAATCAAAACCTGGGATCGGTCCTTGGG 1916
QY 1729 GAAGTGTGCTGCTTGTAAAGCTGACACTCAGCTCTGTAGCCCTGAAGACTCTCAGG 1788
DB 1917 GAAGTGTGCTGCTTGTAAAGCTGACACTCAGCTCTGTAGCCCTGAAGACTCTCAGG 1976
QY 1789 AGAAGGATGCTCTGAACCGGAATCAAGTGGCCCTGTCAAGCTGAGAGGACATCTCTG 1848
DB 1977 AAAAAGGATGCTCTGAACCGGAATCAGGTGGCCACGCTCAAGGCGGAGGACATCTCTG 2036
QY 1849 GCTGAAGCAGACAATGAGTGGGTGCTCAAACTCTACTACTCTCTTCCAGGACAAAGCAGC 1908
DB 2037 GCGGAGGACACAATGAGTGGGTGCTCAAACTCTACTACTCTCTTCCAGGACAAAGCAGC 2096
QY 1909 CTGTACTTGTGATGACTACATACAGGCGGGGATATGATGAGCTCTCTGATCAGGATG 1968
DB 2097 CTGTACTTGTGATGACTACATCTCTGTTGGGACATGATGAGCTCTCTGATCCGGATG 2156
QY 1969 GAGTCTTCCCTGACACCTGCGCGCTTCTACATGTCAGAGTGAACCTGGCCATTGAA 2028
DB 2157 GAGTCTTCCCTGACACCTGCGCGCTTCTACATGTCAGAGTGAACCTGGCCATTGAA 2216
QY 2029 AGTGTCCACAAGATGGGCTTATCCACGGGACATCAAGCTCTGACATCACTATCATGAC 2088
DB 2217 AGTGTCCACAAGATGGGCTTATCCACGGGACATCAAGCTCTGATCACTATCATGAT 2276
QY 2089 CTGATGTCATATTAAGCTGACAGATTTTGGCCTCTGCACTGATTCAGTGGACTCAC 2148
DB 2277 CTGATGTCATATTAAGCTGACAGATTTTGGCCTCTGCACTGATTCAGTGGACTCAC 2236
QY 2149 AATTCCAAGTACTACCAAGAGGACACATGATGACAGACAGCATGGAGCCCGGTGAC 2208
DB 2337 AATTCCAAGTACTACCAAGAGGACACATGATGACAGACAGCATGGAGCCCGGTGAC 2396
QY 2209 CTCTGGGACGATTTTCCAACTGTCGTGTGGACAGGTTAAAGACCTTGGAGCAGAGG 2268
DB 2397 CTCTGGGACGATTTTCCAACTGTCGTGTGGACAGGTTAAAGACCTTGGAGCAGAGG 2456
QY 2269 GCGCAGAGCAGCAGAGGTCCTGGACATTTCTTGTGGGACACCAATATCATC 2328
DB 2457 GCGGGAAGCAGCAGAGGTCCTGGACATTTCTTGTGGGACACCAATATCATC 2516
QY 2329 GCTCCGGAGGTGCTTCTCCGCAAGGTTACACGACGCTCTGACGTGGTGAGCGTCTG 2388
DB 2517 GCACCCGAGGTGCTTCTCCGCAAGGTTACACGACGCTCTGACGTGGTGAGGTGGA 2576
QY 2389 GTGATTCCTTTGAGATGCTGTTGGGACGCCCTTCTTGGCCCGCCACCCACAGAG 2448
DB 2577 GTGATTCCTTTGAGATGCTGTTGGGACGCCCTTCTTGGCCCGCCACCCACAGAG 2636
QY 2449 ACCAGCTGAAGTGTATCACTCGGAGACGACCTGCATATCCCTACGAGGTGAGGCTC 2508
DB 2637 ACCAGCTGAAGTGTATCACTCGGAGACGACCTGCATATCCCTACGAGGTGAGGCTC 2696
QY 2509 AGCGCTGAGCCCGGAGACTCTATCAGAAAGTGTGCTGCGCGGCTGACTGCGCGTGGC 2568
DB 2697 AGCGCTGAGCCCGGAGACTCTATCAGAAAGTGTGCTGCTCGGACACCCGCTGGG 2756
QY 2569 AGGATGGGCGAGATGACCTCAAGGACACCGCTTCTTCAACACCATGCACTTTTCCCGT 2628
DB 2757 CGAAATGGGCGGAGATGACCTCAAGGACACCGCTTCTTCAACACCATGCACTTTTCCCGT 2816
QY 2629 GACATCCGAAGAGGCTGACCTACGTCCTGCTCCCATCAGCCACCCATGACACCTCC 2688
DB 2817 GACATCCGAAGAGGCTGACCTACGTCCTGCTCCCATCAGCCACCCATGACACCTCC 2876

QY 2689 AATTTTGACCGGTGGATGAAGAAAGCCCTTGGCACAGGCGCCAGGAGAGCGGCAAG 2748
DB 2877 AATTTGACCGGTGGATGAAGAAAGCCCTTGGAACAGTGCAGGAGGTAGCACCAAG 2936
QY 2749 GCCTGGGACACGCTGCGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTC 2808
DB 2937 GCCTGGGACACGCTGCGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTC 2996
QY 2809 ACCTTCCGACAGTCTTGTGATGACAAAGGCTATCCCTTCCGGTCCCGGAGGCGCTCAG-- 2866
DB 2997 ACCTTCCGACAGTCTTGTGATGACAAAGGCTATCCCTTCCGGTCCCGGAGGCGCTCAG-- 3056
QY 2867 -----AGCCCGCAGAGAGTGCAGACCCAGGAGTGGGACTTGGAGGTGGGCGC 2916
DB 3057 GCAGAGCTTCCAGAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3116
QY 2917 GAGGCTGCCAGCGGCTGACGTGTA 2942
DB 3117 GAGGCTGCCAGCGCTGACGTGTA 3142

RESULT 7
AR201456
LOCUS AR201456 3213 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6359193.
ACCESSION AR201456
VERSION AR201456.1 GI:20252344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 3213)
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
FEATURES
Location/Qualifiers
1..3213
Source /organism="unknown"
BASE COUNT 946 a 764 c 696 g 807 t
ORIGIN

Query Match 25.8%; Score 812.6; DB 6; Length 3213;
Best Local Similarity 65.1%; Pred. No. 5.5e-170;
Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTCACTGCGCCGCCACACAGGTCACCGCGTACGCGCCGACACATC 1044
DB 919 CACTCTGCTAATTCAGGCTTCTGCCACTACAGTCACTGCCATCACACCCGCTCTATT 978
QY 1045 CTTACCCCTGTGAAGAGCGTGTGTGTCGCGCCCGGAGCCCGCAGACAGCCGTGGGCCC 1104
DB 979 CAACAGCCGCTCAAAAGCATGCGCTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCA 1038
QY 1105 TCGCACCCCGCTGGTGGCTGCGGCCACAGACACTGCCACTGAGAGCTGGAGCAGAG 1164
DB 1039 ACCCATCTCTTGTGATGCCAGCCAGTTCAGACTGTTTCAGCCCTTCTTCTCAG 1098
QY 1165 GAGGCGCAGCGAGGCCACACCCGCTGGATGTGGACTATGGGCTCCGAGCGAGGTG 1224
DB 1099 GGTACAGCTTCAAGTGTGCTGTATCCCACTGTGTGCTGAAGTCCAGGTATCAAGGT 1158
QY 1225 CCACCGCTTCGATTCACAAAGCACTTGTGTCGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
DB 1159 CCACCGCTTCGATTCACAAAGCACTGTCTACACCAAAACCCATCTGTCCTCCATATGAG 1218
QY 1285 GACCTGGACAGCGCTGTCACCAAGTGTGCAGCAGAGTGTGCGAGGGGCGACTGATCTAGAC 1344
DB 1219 TCAGTAAGTAAAGCTGTCAGAGAGTGAACAGCTTACCCAAAGAGGATGATGATGAG 1278
QY 1345 GGGAGTGAACAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGGCAGAGACAAAGCAG 1404
DB 1279 AAGAGTGGGACAGT-----GGTCACTCTGGGATAAAGAAAGAACAG 1323

QY 1405 ATTACAGCTCCCGGCTGCTTCCGCAAGAATACGAGATGAAGAGAGAGAGTCT 1464
Db 1324 ATTACAACTTCCCTATCACTGTCGGAAGAAACAGAAAGAGTGAAGAGAGTCT 1383
QY 1465 CGCATCAAGCTTACTCCCTTATCCCTTCAAAATCTTCTCATGAGCAACAGCTGAGAAAT 1524
Db 1384 CGGATTCAGCTTACTCCCGCAGGCTTTAAGTCTTCTCATGAGCAGCAGCTAGAGAAC 1443
QY 1525 GTTCATCAAAACCTACACAGAGAGTTCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCC 1584
Db 1444 GTCTGAAGTCTCATCAGCAGGCTGTCATCGAAGAACAGCTAGAAATGAATGATG 1503
QY 1585 AAAGCTGGCTCTGTGAGCGCAGCAGAGAGATGAGGAAGATCCTTACCAAGAGGAG 1644
Db 1504 CGGGTTGGATTATCTCAAGATCCAGGATCAATGAGAAGATGCTTTGCCAGAAAGAG 1563
QY 1645 TCTAACTACACCGCTGAAGAGGCCAAGATGGACAAGTCTATGTTGTGAATCAAG 1704
Db 1564 TCTAACTATATTGCTCTTAAAGGGCTAAATGGACAAGTCTATGTTGTGAAGATAAG 1623
QY 1705 ACTTAGGCATCGGTGCCCTTTGGGAAGTGTGCCCTGTGTAAGCTGGACACTCAGCT 1764
Db 1624 ACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCAATCAAGCT 1683
QY 1765 CTGTAGGCATGAAGACTCTCAGGAAGAGAGTGTCTTGAACCGGAATCAAGTGGCCCAT 1824
Db 1684 TTGTATGCAACAAAGACTCTTCCGAAAGAGAGCTTCTGCTCCGAAATCAGGTGCTCAT 1743
QY 1825 GTCAAGGCTGAGGAGACATCTTGGCTGAAGCAGACAAATGAGTGGTGTCAAACTCTAC 1884
Db 1744 GTGAAGCGGAGAGGAGATCTCTAGCAGAGCGACAAATGAGTGGTGTCCGCTGTAC 1803
QY 1885 TACTCTCTCCAGCAAGGACAGCTGTACTTTGTGATGAGTACATACCGCGGGGAT 1944
Db 1804 TACTCTCTCCAGCAAGGACAGCTGTACTTTGTGATGAGTACATCTCTGGGGGAT 1863
QY 1945 ATGATGAGCTCTGATCAGGATGAGGTCTTCCCTGAGCACCTGCCCGCTTCTACAT 2004
Db 1864 ATGATGAGCTTATTAATAGATGGCATCTTCCCTGAAATCTGCGCAGATCTACATA 1923
QY 2005 GCAGAGTGTACCTGGCCATCAAGATGTCACAGATGGGCTTTATCCACCGGACATC 2064
Db 1924 GCAGAGTGTACCTGGCAGTGAAGATGTCATAAATGGGTTTATTCATAGAGATAT 1983
QY 2065 AAGCTGTCAACATCTATCATGACCTGGATGCTATATTAAGCTGACAGATTTGGCCCT 2124
Db 1984 AAGCTGTAAACATTTGATGACCGTGTAGCCATATTAATGACTGACTTTGGCTTG 2043
QY 2125 TCACTGATTCAGTGGACTCACAATTCACAGTACTACCAAGAGGAAACCAATGAGA 2184
Db 2044 TCACTGCTTTCAGATGGACATGACTCCTCAAGTACTACAGAGTGGGATCACCACCG 2103
QY 2185 CAGGACAGATGGACCGGCTGACCTCTGGGACGATGTTTCAACTGCTGCTGGAGAC 2244
Db 2104 CAAGATGATGATTTTCAGTAAAGATGGGAGATCTTCAATTTGCTGGTGGGAC 2163
QY 2245 AGGTTAAAGACCTCGAGCAGAGGCGGAGAGCAGCAGCAGCTGCTGCGCATTTCT 2304
Db 2164 AGACTGAAGCCTAGGAGGAGAGCTGCTCCGAGCAGCAGCTGCTACCCATCT 2223
QY 2305 CTGTGCGGACACCAATTTACATGCTCCGAGGCTGCTTCCGCAAGAGGTTACAGCAG 2364
Db 2224 CTGGTTGGGACTCCCAATATATTGACCTGAAGTGTCTACTGCAACAGGATATACAG 2283
QY 2365 CTCTGACTGTGAGGCTCGGTGTGATCTCTTTGAGATCTGTTGGGAGCGGCT 2424
Db 2284 CTGTGTGACTGTGAGGTTGGTGTATTTCTTTGGAATGTTGCTGGGAGCAACTCT 2343
QY 2425 TTCTTTGGCCCCCACCACCAAGAGAGCTGAGTGAAGTGTCACTGAGGAGCAGCTG 2484
Db 2344 TTCTTTGGCAACACCCCATAGAACACAAATGAAGTTATCATCTGGCAAACTCTCTA 2403
QY 2485 CATATCCCTACGAGGTGAGGCTCAGCGCTGAGCGCCGAGACCTCATCAGGAAGCTGTC 2544

Db 2404 CACATCCCTCTCAAGCTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAACTGTCT 2463
QY 2545 TCGCGGCTGACTCGCGCTGGCGGAGGATGGGAGATGACTCAAGCAGCACCCTGTC 2604
Db 2464 CGAGGACCAAGAGACCGCTCGGCAAGACGGTGTGATGAGATAAGGCTCATCATTT 2523
QY 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGCAGGCTGCACCTTACCTCCAC 2664
Db 2524 TTTAAGACCATGATTTCTTAGTATCTGAGACAGCAGCTGCTTTCATACATCCCTAA 2583
QY 2665 ATCAGCACCCCATGAGACACTTCAATTTTGTACCCGCTGGATGAAGAAAGCCCTGGC 2724
Db 2584 ATCAGCATCCCAACAGATACATCAATTTTCGACCTTTCGAGGTTTTTGTGACAAATG 2643
QY 2725 GAGGACGCGGAGAG---AGCGCCAGGCTGGGACAGCTGGCTCCCGCAGCAGCAG 2781
Db 2644 GATGCGAGGAGGAGAAATATCAGTCTGAGCGGATGATATAAATGGGAAG 2703
QY 2782 CATCAGAGCAGCTTCTTATGAGTTCACCTTCCGAGGTTCTTCGATGACACAGGCTAT 2841
Db 2704 CACCCGAGCAGCTTCTTATGAGTTCACCTTTCGAGGTTTTTGTGACAAATGCTAC 2763
QY 2842 CCTTCCGCTGCCGAAAGCCCTCAGAGCCGCGCAGAGAGT 2880
Db 2764 CCATATAATTATCCAAAGCCTATTGAGTATGAATACATT 2802

RESULT 8
AF104414
LOCUS AF104414 3213 bp mRNA linear ROD 18-MAR-1999
DEFINITION Mus musculus large tumor suppressor 1 (Lats1) mRNA, partial cds.
ACCESSION AF104414
VERSION AF104414.1 GI:4324435
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 3213)
AUTHORS Xu,T., Wang,W., Zhang,S., Stewart,R.A. and Yu,W.
TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase
JOURNAL Development 121 (4), 1053-1063 (1995)
MEDLINE 95262551
PUBMED 7743921
REFERENCE 2 (bases 1 to 3213)
AUTHORS Tao,W., Zhang,S., Turenchalk,G.S., Stewart,R.A., St John,M.A., Chen,W. and Xu,T.
TITLE Human homologue of the Drosophila melanogaster lats tumour suppressor modulates CDC2 activity
JOURNAL Nat. Genet. 21 (2), 177-181 (1999)
MEDLINE 99440767
PUBMED 9988268
REFERENCE 3 (bases 1 to 3213)
AUTHORS St John,M.A., Tao,W., Fei,X., Fukumoto,R., Carcangiu,M.L., Brownstein,D.G., Parlow,A.F., McGrath,J. and Xu,T.
TITLE Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction
JOURNAL Nat. Genet. 21 (2), 182-186 (1999)
MEDLINE 99440768
PUBMED 9988269
REFERENCE 4 (bases 1 to 3213)
AUTHORS Tao,W. and Xu,T.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1998) Genetics, HHMI, Yale University, 295 Congress Ave BCM 254D, New Haven, CT 06536, USA
FEATURES
Location/Qualifiers
1. 3213
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="newborn"

Query Match	25.8%	Score 812.6	DB 10	Length 3213
Best Local Similarity	65.18	Pred. No. 5.5e-170		
Matches 1237	Conservative 0	Mismatches 644	Indels 18	Gaps 2
Qy	985	CAACCTGAGCCCTCACTGCGCGCCGCCAACAGGTCACCGCGCTGAGCGCGCCACACATC	1044	
Db	919	CACCTGCTAAATTCAGCGCTTCTCGCTACTACAGTACTGCCATCACACCGCTCTCTATT	978	
Qy	1045	CTTTCACCTGTGAAGAGCGTGGTGTCTGGGGCCGAGCGCCAGACCGCTGGGGGCC	1104	
Db	979	CACAGCCCGTGAAGCATGCGCTCTGAAACAGAGCTGCAGACTGCTTTAGGCCCA	1038	
Qy	1105	TGCGACCCCGCTGGTGCTGGCCCCACAGACCTGCCACTGAGAGCCTGGAGACGAAG	1164	
Db	1039	ACCATCTCTTCTGGATGCGCAGCGAGTTCAGACTGTTACGCTACCCCTTTTCTGAG	1098	
Qy	1165	GAGGGCAGCGAGCCGCACACCCGCTGGATGPGGACTATGGCGGTTCGAGCGCAGGTGC	1224	
Db	1099	GGTACAGCTTCAAGTGTGCTGTCATCCACACCTGTTGCTGAAGCTCCAAGCTATCAAGGT	1158	
Qy	1225	CCACCGCTCCGTATCCAAAGCACTTGCTGCTGCCAGTAAGTCTTGAGCAGTACAGCGTG	1284	
Db	1159	CCACCGCGGCTTATCCAAAGCATCTGCTACACCAAAACCCATCTGTCCCTCCATATGAG	1218	
Qy	1285	GACCTGGACAGCCTGTGSCACCACTGTGCAGCAGAGTGTGCGAGGGGCGACTGTACTAGAC	1344	
Db	1219	TCAGTAAGTAAAGCCCTGCAAGATGAACAGCCTAGCTTACCCAAAGGAAGATGAATGTAG	1278	
Qy	1345	GGGAGTACAGAGCCCAAAAGGTGCGAAGGGAGACAAAGCTGGCAGAGACAAAAGCAG	1404	
Db	1279	AAGAGTGGGHCAGT-----GGTGACTCTGGGATAAAGAAAGAAACAG	1323	
Qy	1405	ATTTCAGACCTCCCGGTGCTGTCCCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT	1464	
Db	1324	ATTCAACATTCACCTATCACTGTTCGGAAGAAACAGAAAGATGAGACGAAAGAGTCT	1383	
Qy	1465	GCATCAAGAGTTACTCCCTTTATGCCCTTCAAAATTTCTCATGGAGCAACACTGGAGAAAT	1524	
Db	1384	CGGATTCAGATTACTCCCAACAGCCCTTTAAGTTCTTCATGGAGCAGCACGTAGAGAAC	1443	
Qy	1525	GTCACTAAAAACCTACACGAGAGAGGTACGCCGGAGGCTACGCTGGAGCAGAAATGCC	1584	
Db	1444	GTCTCGAAGTCTCATCAGCAGCGGCTGCTGCATCGGAAGACGAGTCAAAATGAATGATG	1503	

Db 2584 ATCACGATCCACACAGATACATCCATTCGACCCCTGTGTATGATGATAAATTTGGAGC 2643
 Qy 2725 GAGGCCAGCGGAGAG---AGGCCAGGCGCTGGGACACCGCTGCCCTCCCGGACGACAG 2781
 Db 2644 GATGGCAGGAGGAGGAGAAATATCAGTGACACTCTGAGCGGATGGGTATATAAATGGGAAG 2703
 Qy 2782 CATCCAGACGACGCTTCTATGAGTTCACTTCCGAGGCTTCTGATGACACAGGCTAT 2841
 Db 2704 CACCCGAGCAGCGCTTCTATGAGTTCACTTCCGAGGCTTCTGATGACACAGGCTAT 2763
 Qy 2842 CCCTCCGCTGGCCGAGGAGCCCTCAGAGCCGCGAGAGGT 2880
 Db 2764 CCATATAATTCACCAAGCGCTATTGAGTATGAATACATT 2802

RESULT 9
 AR201455
 LOCUS AR201455 3984 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 3 from patent US 6359193.
 ACCESSION AR201455
 VERSION AR201455.1 GI:20252343
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3984)
 AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
 TITLE Nucleotide sequences of lats genes
 JOURNAL Patent: US 6359193-A 3 19-MAR-2002;
 FEATURES
 Location/Qualifiers
 1..3984
 /organism="unknown"

BASE COUNT 1280 a 847 c 798 g 1059 t
 ORIGIN

Query Match 24.1%; Score 760; DB 6; Length 3984;
 Best Local Similarity 63.7%; Pred. No. 2.7e-158;
 Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;
 Qy 985 CAACCTGAGCCCTCACTGCGCCGCCCAACACGCTCAGCGCGTGAAGCGCCGACACATC 1044
 Db 1653 CACTCTGCTAAATTCAGCGCTTCTCTACACAGCTCACTGCAATTACACGACTCCTATT 1712
 Qy 1045 CTTTCACTGCTGAAGAGCGTGTGCTGCGGCCGAGCCGACAGACCGCTGGGCCC 1104
 Db 1713 CAACAGCTGTGAAGATGCTGTATTAAACAGAGCTACAGACTGCTTTAGCACT 1772
 Qy 1105 TCGACCGCGCTGGTGGCTGCGCCACAGACCTGCCACTGAGAGCGCTGGAGCGAG 1164
 Db 1773 ACACACCGCTTCTGGATACACAGCCAAATTCAAACTGTTCACCCAGTCTCTTCTGAG 1832
 Qy 1165 GAGGCAGCGGAGCCACACCGCTGATGTGGACTATGCGCGCTCCGAGCGCAGGTGC 1224
 Db 1833 GGAACCGCTTCAATGTGACTGTGATGCCACTGTGTGCTGAAGCTCCAACTATCAAGA 1892
 Qy 1225 CCACCGCTTCCGTATCCAAAGCACTGTGCTGCCAGTGAAGTCTGAGCAGTACAGCGTG 1284
 Db 1893 CCACACACCGCTACCCAAAGCACTGTGCTGACAAACCCATCTGTTCTCCATACAG 1952
 Qy 1285 GACCTGGAGAGCGCTGTGACACAGTGTGACAGAGTGTGCGAGGGGCGCTGATAGAC 1344
 Db 1953 TCAATCAGTAAGCTTAGCAAGAGGATCAGCCAAAGCTTGCCCAAGGAAGTCAAGTGA 2012
 Qy 1345 GGGAGTGACAGAGCCACAAAGTGGGAGGAGACAAAGCTGGGAGAGCAAAAAGCAG 1404
 Db 2013 AAGAGTTA-----TGAAATGTGTAGTGGGATATAAAGAAAGAAACAG 2057
 Qy 1405 ATTCAAGCTCCCGGTGCTCTCGCAAGATAGCAGAGTGAAGAGAGAGAGTCT 1464
 Db 2058 ATTCAAGCTTCACTTACTGTTAGGAAAACAGAAAGTGAAGAGGAGGAATCT 2117
 Qy 1465 CGCATCAAGAGTTACTCCCGCTTATGCTTCAAAATCTTCATGGAGCAACACGTTGAGAT 1524

Db 2118 CGTATTCAAAGTTATTTCTCTCAAGCATTTAAATCTTTTATGGACACATCTAGAAAT 2177
 Qy 1525 GTCAATCAAAACCTACAGCAGAGGTCACCGGAGGCTACAGCTGGAGCAGAAATGGCC 1584
 Db 2178 GTACTCAAAATCTCAGCAGAGGCTTACATCGTAAAAAACAATTAGAAGTAATGAATG 2237
 Qy 1585 AAAGCTGGCTCTGTGAGCGCGAGGAGCAGATGAGGAAGATCTCTTACCAGAGGAG 1644
 Db 2238 CGGTTGGATTATCTCAAGATCCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 2297
 Qy 1645 TCTAACTACAACCGCTGAAGAGGCGCAAGATGGACAGTCCATCTTTGTGAAAAATCAAG 1704
 Db 2298 TCTAATTACATCGTCTTAAAAGGCTAAAATGGACAAGTCTATGTTTGTGAAGATAAG 2357
 Qy 1705 ACTCTAGGATCGGTGCGCTTTGGGGAAGTGTGCTGCTGTTGAAGTGGACACTCACGT 1764
 Db 2358 ACACCTAGGAATAGGAGCATTTGGTGAAGTCTCTTAGCAAGAAAGTAGATCTAAGGCT 2417
 Qy 1765 CTGTAGCGCATGAAGACTCTCAGGAAGAGGATGCTTGAACCGGAATCAAGTGGCCAT 1824
 Db 2418 TTGTATGCAACAAAACCTCTCGAAGAAAGATGTTCTTTCGAAATCAAGTCCGCTAT 2477
 Qy 1825 GTCAAGGCTGAGAGGACATCTGCTGAAGCAGACAAATGAGTGGTGGTCAAACTCTAC 1884
 Db 2478 GTTAAAGGCTGAGAGATATCTGCTGAAGCTGACAATGAATGGTAGTCTGCTATAT 2537
 Qy 1885 TACTCTTCCAGGACAGAGACCGCTGTACTTTGTGATGAGTACATACAGCGGGGAT 1944
 Db 2538 TATTCATTCCAAGATAAGGACAAATTTATCTTTGTAAGTGGACTACATCTCTGGGGT 2597
 Qy 1945 ATGATGAGCTCTGATCAGGATGGAGTCTTCCCTGAGCAGCTGCGCCGCTTCTACATT 2004
 Db 2598 ATGATGAGCTTATTAATGAGATGGCATCTTTCAGAAAGTCTGGCAGCATCTACATA 2657
 Qy 2005 GCAGAGTTGACCTTGCCCATTCGAAAGTGTCCAAAGTGGTGGTGGTGGTGGTGGTGG 2064
 Db 2658 GCAGAACTTACCTGTGAGTTCGAAAGTGTTCATAAAATGGGTTTATTATCATAGATA 2717
 Qy 2065 AGCCTGACACAACTACTCATCAGCTGGATGTCATATTAAGTGGTGGTGGTGGTGGT 2124
 Db 2718 AAACCTGATAATTTTGTGATGCTGATGCTATATAAATGACTGACTTTGGCCTC 2777
 Qy 2125 TGCACCTGATTCAGGTGGACTCACAATTCGAAGTACTACAGAAAGGAAACACATGAGA 2184
 Db 2778 TGCACCTGCTTCAAGTGGACACAGCTTCTAAGTACTATCAGAGTGGTGGTGGTGG 2837
 Qy 2185 CAGGACAGCTGGAGCGCGGTGACCTCTGGGAGAGTGTTCACACTGTGCTGGTGGAG 2244
 Db 2838 CAAGATAGCATGGATTCAGTAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 2897
 Qy 2245 AGTTAAAGACCTGGAGCAGAGGCGGAGAGAGCAGCAGCAGGCTGGCTGGCAGATCT 2304
 Db 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCAGCAGGATGTCTAGCAGATCT 2957
 Qy 2305 CTGTGCGGACACCAAAATACATCGCTCGGAGGTCTTCTCCGAAAGGGTACAGCAG 2364
 Db 2958 TTGGTTGGAGCTCCCAATTAATGACCTGAAGTGTGCTACGAGACAGGATACACAG 3017
 Qy 2365 CTCTGTGACTGTGAGCGGTGCTGTGATCTCTTTGAGATGCTGTTGGGAGCGCGCT 2424
 Db 3018 TTGTGTGATTTGGTGGAGTGTGTTGTTTATTTTAAATGTTGTTGGGAGCAACTCT 3077
 Qy 2425 TTCTTTGGCCCCCCCCCAGCAGCAGAGCGAGCTGAAGTGAATCACTGGGAGAGCAGCTG 2484
 Db 3078 TTCTTTGGCAAAACACCAATTAGAAAACAAATGAAGGTATCAACTGGCAAACTCTCT 3137
 Qy 2485 CATATCCCTACGAGCTGAGGCTCAGCGCTGAGGCGCGGAGCGCTCATCAGAGGTGTC 2544
 Db 3138 CACATTCCACCAAGCTTAACTCAGTCTGAGCTCTGATCTTATTTATTTAACTTTGC 3197
 Qy 2545 TCGCGGCTGACTGCGCTGGGCGGATGGGCGAGATGACCTCAAGSCACACCGCTTC 2604

```

Db 3198 CGAGGACCCGAGATCGCTTAGGCAAGAAATGCTGCTGATGAATAAAGAGCTCATCCATTT 3257
Qy 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAGCAGGCTGCACCCCTACGTCCTCCACC 2664
Db 3258 TTTAAACAATGACTTCTCCAGTACACCTGAGACAGACAGTCTGCTTCATACATCTCTAAA 3317
Qy 2665 ATCAGCACCCATCGGACACCTTCAATTTTACCGGTGGATGAAGAAGCCCTGGCAC 2724
Db 3318 ATCACACACCAACAGATACATCAAAATTTTATCTCTGATCTCTGATGAATAATGAGAGT 3377
Qy 2725 GAGGCCAGC---GGAGAGAGGCGCAAGCGCTGGACAGCGCTGGCTCCCGCAGCAGCAAG 2781
Db 3378 GATGATACGAGGAGAAATGTAATGACACTCTCAATGATGATGATAAATAATGGAAG 3437
Qy 2782 CATCAGACGACGCTTCTATGAGTTTCACTTCCGCGAGTTCTTCATGACACACCGGCTAT 2841
Db 3438 CATCTGAACATCATCTATGAATTTACCTTCGGAAGTTTTCATGATGACAAATGGCTAC 3497
Qy 2842 CCCTTCGGTCCCGAAGCC 2861
Db 3498 CCATATAATATTCGGAAGCC 3517

RESULT 10
AF164041
LOCUS Homo sapiens WARTS protein kinase (WARTS) mRNA linear PRI 06-DEC-1999
DEFINITION AF164041
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4241)
AUTHORS Kishimoto,T., Niwa,S.-i., Nagamine,Y., Nishiyama,Y. and Saya,H.
TITLE WARTS protein, polynucleotide encoding the same, antisense
polynucleotide thereof, and antibody recognizing the protein
JOURNAL Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-Ku,
Yokohama-shi; Kanagawa;
Japan;
REFERENCE 2 (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiraoka,T., Kitamura,N. and
Saya,H.
TITLE A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
JOURNAL FEBS Lett. 459 (2), 159-165 (1999)
MEDLINE 99447636
PUBMED 10518011
REFERENCE 3 (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto 860-0811,
Japan

FEATURES
source Location/Qualifiers
1. .4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .4241
/genes="WARTS"
CDS 33..3425
/genes="WARTS"
/notes="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"

```

```

/db_xref="GI:5738136"
/translation="MKRSEPEGYRQMRPKTFPASNTVTVSSRMLQIRESLRNLSP
SDAAKAEHNMSKSTEDPROVRNPPKFGTHKALQELRNSLLPFANETSSRSTSEVN
POMLODLOAAGFEDDMVIOALOTNNRSIEAAIEFISKMSYQDPREBOMAAAAARPIN
ASKMKGNOQSVNRKSKESLVPORHGPPLGESVAYHSESPNSQTDVGPPLSGS
GISAFVQAHFNSGQRVNPPTPPQVRSVTPPPRGGTTPPRGTTTPPPSWEPNSQTKR
YSKMEYVTSRISPSVPGTGWQEGYPPPLNTSPMNPNGOGRGISSVPGVGRQPIIMOS
SSKFNFPGRPGNGQTDFMIHQNVAGTVNPPPPYPLTAANGQSPALQIG
GSAAPSNTNGSIPQSMVPRNHNHMLYINISVGLQTNWQSSAPASQSPSSGHE
IPTWOPNIPIVRSFNPLNASHANSOPSATVTITPAIQQPVKSMRVLLKPEL
OTALATHPSKIPQIQTQVQSPPECTASNTVMPVPAEAPNYOQPPPPYKXHLHQ
NPSPYPIESIKSKEDQPSLPKEDSEKSENVDSGDKKKQITTSPIIVRNKKDE
ERESRQISVPOAFKFMHQHVENLKHQRLKRRKOLENEMQITRTALYATKLRKD
KMLQKESNYIRLKRQKMSFVKIKTIGIFAGFEYCLARKYDTKALYATKLRKD
VLLRNOVHVKAERDILAEADNEMVRLYISFQKONLYFVMDYIPGDMNSLLIRMG
IFPESLARFYIAELTCAVESVHKMGFTHRDIKPDNLIKIDRGHILKIDFGCTGFRWT
HDSKYVSGDHPQDSMDFSNEMGDPSSCRGDLKPLERARAQRORCLAHSLGTPP
NYTAPEVLLRTGTQLCDWWSGVILFEMLVGQPPFLAQTPLETOMKVINWQISLHIP
POAKLSPEASDLIIKICRGEDRLGNKNGADEIKAHFFFTKIDFSSDLRQOSASYPKI
THPTDTSNFDVDPDKLWSDNNEENVNDTLNGVYKNGKHPHEFAFEFTFRFDDNG
YPYNYPKPIEYINSQGSBQSDDEDDQNTGSEIKNRDLVYV"
misc_feature 2145..3059
/genes="WARTS"
BASE COUNT 1384 a 884 c 817 g 1156 t
/notes="Region: serine/threonine kinase domain"
Query Match 24.1%; Score 760; DB 9; Length 4241;
Best Local Similarity 63.7%; Pred. No. 2.8e-158;
Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;
Qy 985 CAACCTGAGCGCTCACATGCGCGCCGCCAACACGCTCACCGCGTGACGGCGCACACATC 1044
Db 1455 CACTCTGCTAATCTCAGCCTTCTGCTTACACAGCTACTGCAATTACACAGCTCTTATT 1514
Qy 1045 CTTGACCTGTGAAGAGCGTGGTGGTGGCGCCGAGAGCCGCGTGGGGGCC 1104
Db 1515 CAACAGCGCTGTGAAGAGTATGCGTGTATTAAACACGAGAGCTACAGACTGCTTTAGCACCT 1574
Qy 1105 TCGCACCCCGCTGGTGGTGGCGCCACACGCTCCACTGACGAGCGCTGGAGACGAAG 1164
Db 1575 ACACACCTCTCTGGATACACAGCCCAATTCAACTGTTCACCCAGTCTCTTCTCTGAG 1634
Qy 1165 GAGGCGAGCGAGCCGACACACCGCTGGATGTGGAGTATGGCGCTCCGAGCGCAGGTGC 1224
Db 1635 GGAACCGCTTCAATATGTGACTGTGATCCACTCTTCTGTAAGCTCCAAACTATCAAGGA 1694
Qy 1225 CCACCGCTCGGTATCAAAGCACTTCTGCTGCCAGTAGTCTGAGCAGTACAGCGTG 1284
Db 1695 CCACCAACCGCTTACCCAAACATCTGCTGCCAAACCCATCTCTCTCTCCATACGAG 1754
Qy 1285 GACCTGCACAGCGCTGTGCACAGTGTGCAGAGTCTGGAGGGGCGCTGATCTAGAC 1344
Db 1755 TCATCATAGTAACTTACAAAGAGGATCATCAAGCTTGCCCAAGGAAGATGAGAGTGAA 1814
Qy 1345 GGGAGTGACAGAGCCACAAAGGTGCGAAGGGAGACAAAGCTGGCAGAGACAAAGACAG 1404
Db 1815 AAGAGTTA-----TGAAATGTTGATAGTGGGATAAAGAAAGAAACAG 1859
Qy 1405 ATTCAGACCTCCCGGTGCTGTCCGCAAGAAATAGCAGAGATGAAGAGAGAGAGAGTCT 1464
Db 1860 ATTACAACTTCACTTACTGTAGGAAAACAAAGAAAGATGAAGAGCGAAGGAATCT 1919
Qy 1465 CGCATCAAGAGTTACTCCCTTATGCTTCAATTTCTTTCATGGAGCAACAGCTGGAGAA 1524
Db 1920 CGTATTCAAAGTTATCTCTCTCAAGCAATTTAAATCTTTATGGAGCAACATGTAGAAAT 1979
Qy 1525 GTCATCAAAACCTACCAAGAGGTGACGGAGGTCTACAGCTGGAGCAAGAAATGCC 1584
Db 1980 GTACTCAAACTCATCAGCAGCGTCTACATCGTAAAAAACAATTAGAGAAATGAATGATG 2039
Qy 1585 AAGCTGGGCTCTGTGAGGCGGAGCAGAGATGAGGAAGATCCTCTCTACCAGAGGAG 1644

```

Db 2040 CGGGTGGATTATCTCAAGATCCCGAGGATCAAAATGAGAAGATGCTTTGGCCAAAAGAA 2099
 QY 1645 TCTAATCAACCGCTGAAGAGGCCAAGATGGACAGCTCCATGTTCTGGAATCAAG 1704
 Db 2100 TCTAATCACTCGCTTAAAGAGGCTAAATGGACAGCTCTATGTTCTGGAAGATAAG 2159
 QY 1705 ACTCTAGGCATCGGTGCCCTTTGGGAAGTGTGCCCTGCTTGAAGGTGGACACTCACGCT 1764
 Db 2160 ACACCTAGGAATAGGAGCAATTTGGTGAAGTCTGTAGCAAGAAAGTAGATACCTAAGGCT 2219
 QY 1765 CTGTAGGCATCAAGACTCTCAGGAAGAGGATGCTCCTCAACCGGAATCAAGTGGCCAT 1824
 Db 2220 TTGTATGCAACAAAACCTTCCGAAGAAAGATGTTCTTCTCGAATCAAGTGGCTAT 2279
 QY 1825 GTCAAGGCTGAGAGGACATCTCGCTGAAGCAGACAAATGAGTGGGTGCTCAACTCTAC 1884
 Db 2280 GTTAGGCTGAGAGATATCTCGCTGAAGCTGACATGAATGAGTGGTAGTCTGCTATAT 2339
 QY 1885 TACTCCTTCCAGGACAAGGACAGCTGTACTTTGTGAGTACTACATACAGGCGGGGAT 1944
 Db 2340 TATTCATTCCAAGATAAGGACAATTTATCTTTGTAATGAGTACTACATTCCTGGGGTGTAT 2399
 QY 1945 ATGATGAGCCTCTGATCAGGATGAGGCTCTTCCCTGAGCAGCTGCGCGCTTCTACATT 2004
 Db 2400 ATGATGAGCCTTATTAATTAAGATGGCATCTTCCAGAAAGTCTGGCAGGATTTACATA 2459
 QY 2005 GCAGAGTTGACCTGGCCATTGAAAGTGTCCACAGATGGGCTTTATCCACCGGACATC 2064
 Db 2460 GCAGAGCTTACCTGTGAGTGTCAAGTGTTCATAAATGGGTTTATTCATAGAGATAT 2519
 QY 2065 AAGCTGTCAACATCTACTCATGACCTGGATGCTCATATTAAGTGTGACAGATTTGGCCCT 2124
 Db 2520 AAACCTGATAATTTATGATGATGCTGATGCTCATATTAATTAATGACTGACTTTGGCCCT 2579
 QY 2125 TGCATGAGTTCAGTGGAGTCAATCCAGTACTACCAAGAGGAGGACACATGAGA 2184
 Db 2580 TGCATGAGTTCAGTGGAGTCAATCCAGTACTACCAAGAGGAGGAGGACACATGAGA 2639
 QY 2185 CAGGACAGATGGAGCCGCTGACCTCTGGGACGATGTTTCCAACTGTGCTGTGGAGAC 2244
 Db 2640 CAAGATAGATGATTTAGTAAATGAATGGGGGATCCCTCAAGCTGTGCTGTGGAGAC 2699
 QY 2245 AGTTTAAAGACCTGGAGAGGCGGAGAGAGCAGCAGGAGTGTGCTGACATTTCT 2304
 Db 2700 AGACTGAAGCCATTAGAGCGGAGAGTGCACCCAGCAGCAGGATGTCTAGCAGATTTCT 2759
 QY 2305 CTGTGCGGACACCAATATACATCGCTCCGAGGTGCTTCTCGGAAAGGGTACAGGAG 2364
 Db 2760 TTGGTTGGGACTCCCAATTTATATTCACCTGAAGTGTGCTACGACAGGATACACAG 2819
 QY 2365 CTCTGTGAGTGTGGAGCTGCTGTGATTTCTTTGAGATGCTGTTGGGAGCGGCT 2424
 Db 2820 TTGTGTGATTTGGAGTGTGGTGTATTTCTTTTGAATTTGGTGGGACAACTCTCT 2879
 QY 2425 TTCTTGGCCCCCACCACACAGAGAGCGAGCTGAAAGTGTATCACTGGGAGAGCAGCTG 2484
 Db 2880 TTCTTGGCAACAAACACCATTTAGAAACACAAATGAAGTTATCAACTGGGCAACATCTCT 2939
 QY 2485 CATATCCCTACAGGTGAGGCTGAGGCTGAGGCTGAGGCGGAGGAGCTCATCAGAGCTGTC 2544
 Db 2940 CATATCCACCAACAGCTAAACTAGCTAGCTGAGGCTTCTGATCTTATTTAAACTTTGC 2999
 QY 2545 TCGCGGCTGACTCGCGCTGGCGAGGATGGGAGATGAGCTCAAGGACACACCGGTC 2604
 Db 3000 CGAGACCGGAGATCGCTTAGCGAAGATGGTGTGATGAATTAAGCTCATCAATTT 3059
 QY 2605 TTCAACACCATCGACTTTTCCGTGACATCCGAAGAGGCTGCACCTTACGTCGCCACC 2664
 Db 3060 TTATAAACAATTTGACTTCTCCAGTGACCTGACGACAGCAGCTGCTTCATACATTTCTTAA 3119
 QY 2665 ATCAGCACCCTCAGACACCTCAATTTTACCCTGGGATGAGAAAGGCCCTTGGCAC 2724
 Db 3120 ATCACACACCCCAACAGATACATCAAAATTTGATCTCTGTTGATCTCTGATTAATTTGGAGT 3179

QY 2725 GAGGCCAGC---GGAGAGAGCCCAAGGCTGGGACACGCTGCCCTCCCCAGCAGCAAG 2781
 Db 3180 GATGATAACGAGAGAAATTAATGACACTCTCAATGGATGATATAAAATGGAAG 3239
 QY 2782 CATCCAGACGACGCTTCTATGAGTTCACCTTCCGACGCTTCTTCGATGACACACGCTAT 2841
 Db 3240 CATCTGAACTGATCTATGATTAATTTACCTTCCGAAGGTTTTTTTATGACAAATGGCTAC 3299
 QY 2842 CCTTCCGGTGGCCGAGCC 2861
 Db 3300 CCATATAATTATCCGAAGCC 3319

RESULT 11
 AF104413
 LOCUS Homo sapiens large tumor suppressor 1 (LATS1) mRNA, complete cds.
 DEFINITION AF104413
 ACCESSION AF104413
 VERSION AF104413.1 GI:4324433
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 4307)
 AUTHORS Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.
 TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila
 lats gene encodes a putative protein kinase
 JOURNAL Development 121 (4), 1053-1063 (1995)
 MEDLINE 9526251
 PUBMED 7743921
 REFERENCE 2 (bases 1 to 4307)
 AUTHORS Tao, W., Zhang, S., Trenchalk, G.S., Stewart, R.A., St John, M.A.,
 Chen, W. and Xu, T.
 TITLE Human homologue of the Drosophila melanogaster lats tumour
 suppressor modulates CDC2 activity
 JOURNAL Nat. Genet. 21 (2), 177-181 (1999)
 MEDLINE 99140767
 PUBMED 9988268
 REFERENCE 3 (bases 1 to 4307)
 AUTHORS St John, M.A., Tao, W., Fei, X., Fukumoto, R., Carcangiu, M.L.,
 Brownstein, D.G., Parlow, A.F., McGrath, J. and Xu, T.
 TITLE Mice deficient of lats1 develop soft-tissue sarcomas, ovarian
 tumours and pituitary dysfunction
 JOURNAL Nat. Genet. 21 (2), 182-186 (1999)
 MEDLINE 99140768
 PUBMED 9988269
 REFERENCE 4 (bases 1 to 4307)
 AUTHORS Tao, W. and Xu, T.
 TITLE Direct Submission
 JOURNAL Submitted (04-NOV-1998) Genetics, HMTI, Yale University, 295
 Congress Ave BCM 254D, New Haven, CT 06536, USA

FEATURES
 source
 1..4307
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="brain"
 /dev_stage="fetus"
 1..4307
 /gene="LATS1"
 /note="Putative serine threonine kinase"
 /codon_start=1
 /product="large tumor suppressor 1"
 /protein_id="AAD16882.1"
 /db_xref="GI:4324434"
 /translation="MKSEKPEGYROMPKTFPASNYTVSSROMLOETRESLRNLSKP
 SDAKAENKMSKSTEDPROVNPKEFGTHHKALEIRNSLLLPANETNSRSTSEVN
 PQMLQDLQAAGFEDMTVQLQNNRSIEAAIEFISKSYQDPPRROMAAAAARPN
 ASMKPGNVQSVNRKQSKESLPQRHGPPLGESVYHSESPNSQTDYGRPLSGS
 GISAFVQAHPSNGQVRNPPPPPPQVRSVTPPPPRGQTTPPPRGTTPPPSWEPNSQTKR

gene
 CDS

Query Match	24.1%	Score 760;	DB 9;	Length 4307;
Best Local Similarity	63.7%;	Pred. No. 2.8e-158;		
Matches 1197;	Conservative -0;	Mismatches 665;	Indels 18;	Gaps 2;
QY	985	CAACCTGAGCCCTCACTGCCGCCGCCCAACACGGTCAACCGCTGAGCGCCGACACATC	1044	
Db	1971	CACCTGCTAAATCTCAGCCTTCTGCTACACAGTCACTGCAATACACAGCTCCTAT	2030	
QY	1045	CTTCAACCTGTGAAGACGCTGGTGTCTGCGGCCGCCAGCCCGACACACGCCGTGGGGCC	1104	
Db	2031	CAACAGCCTGTGAAGAATATGCGTGTATTAACACAGAGCTACACACTGCTTTAGCACT	2090	
QY	1105	TGCGACCCCGCTGGTGGCTCGGCCACAGCACTGCCACTGAGAGCCTGGAGACGAG	1164	
Db	2091	ACACACCTTCTTGATACACAGCAATTCACAACTGTTCAACCCAGTCCCTTTCTCTGAG	2150	
QY	1165	GAGGCGAGCGCAGGCCACACCCGCTGGATGTGGACTATGGCGCTCGAGCGCAGTGC	1224	
Db	2151	CGAACCGCTTCAAAATGTGACTGTGATGCCACTGTGTGAAGCTTCAAACTATCAAGGA	2210	
QY	1225	CCACCGCTCGGTATCCAAAGCACTTGTGTGCCCACTAAAGTCTGAGCAGTACAGCGTG	1284	
Db	2211	CCACCAACCACTTACCCAAAACATCTGTGCACCAAAACCCTCTGTCTCTCCATACGAG	2270	
QY	1285	GACCTGGACAGCCTGTGCACCACTGTGCAGCAGAGTCTGCCAGGGGCGCACTGTACTAGAC	1344	
Db	2271	TCAATCAGTAAAGCTAGCAAGAGGATCAGCCAGCTGTGCCCAAGAGATGAGAGTGAA	2330	
QY	1345	GGGAGTGACAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGCCAGACACAAAAGCAG	1404	
Db	2331	AAGAGTTA-----TGAAAATGTTGATAGTGGGATAAAGAAACAAACAG	2375	
QY	1405	ATTCAGACCTCCCGGTCCTGTCCGCAAGNATAGCAGAGATGAAGAGAGAGAGTCT	1464	
Db	2376	ATTACAACTTCACTTATCTCTCAAGCATTAATCTTTATGGAGCAACATGTAGAAAT	2435	
QY	1465	CGCATCAAGAGTTACTCCCTTATGCCCTTCAAAATCTTATGGAGCAACAGCTGGAGAAT	1524	
Db	2436	CGTATTCAAAGTTATCTCTCAAGCATTAATCTTTATGGAGCAACATGTAGAAAT	2495	
QY	1525	GTCATCAAAACCTACACAGCAGAGGTCAGCCGGAGGCTACAGCTGGAGCAGGAATGGCC	1584	
Db	2496	GTACTCAAAATCTCATCAGCAGCGCTCTACATCGTAAAAACAAATTAGAGAATGAATGATG	2555	
QY	1585	AAAGCTGGGCTCTGTAGGCCGAGCAGGACAGTACAGTGGAGGATCTCTTACCAAGGAG	1644	
Db	2556	CGGTTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGTGTTGCCAAAAGAA	2615	
QY	1645	TCTAACTAACCGGCTGAAGAGGGCCAGATGGACAAGTCCATGTTGTGAAAAATCAAG	1704	
Db	2616	TCTAAATACATCCGCTTTAAAGGGCTAAAAATGGACAAGTCTATGTTGTGAGATAAAG	2675	
QY	1705	ACTCTAGGCATCGCTTTTGGGGAAGTGTGCCTCGCTTGTAGCTGGACACTACGCT	1764	
Db	2676	ACACTAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAAAGTACATACTAGGCT	2735	

QY 2842 CCCTTCGGTGGCCGAAGCC 2861
Db 3816 CCATATAATTATCCGAAGCC 3835

RESULT 12
E24613 2442 bp DNA linear PAT 18-JUN-2001
DEFINITION warts protein, polynucleotide encoding the protein, antisense
polynucleotide thereof and antibody recognizing the protein.
ACCESSION E24613
VERSION E24613.1 GI:13024646
KEYWORDS JP 1999089580-A/1.
SOURCE unidentified.
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2442)
AUTHORS Toshihiko K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
TITLE warts protein, polynucleotide encoding the protein, antisense
polynucleotide thereof and antibody recognizing the protein
JOURNAL Patent: JP 1999089580-A 1 06-APR-1999;
SUMITOMO ELECTRIC IND LTD

COMMENT OS Unidentified
PN JP 1999089580-A/1
PD 06-APR-1999
PF 24-SEP-1997 JP 1997258689
PR
PI TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
YASUYUKI NISHIYAMA,
PI HIDEYUKI SAYA
PC C12N15/09, C07K14/435, C07K16/18, C1201/68, G01N33/53, PC
G01N33/532, C12P21/02,
PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
(C12N15/00, PC C12R1:91)
CC Strandedness: Double;
CT Topology: Linear;
FH Key
FT source
FT Location/Qualifiers
1..2442
/organism="unidentified"
/db_xref="taxon:32644"

BASE COUNT 777 a 528 c 491 g 646 t

ORIGIN

Query Match 23.7%; Score 747.8; DB 6; Length 2442;
Best Local Similarity 63.5%; Pred. No. 1.4e-155;
Matches 1184; Conservative 0; Mismatches 562; Indels 19; Gaps 2;

QY 985 CAACCTGAGCCCTCACTGCGCCGCCACACGCTCACGCGCTGACGCGCCGACACATC 1044
Db 592 CACTCTGCTAATCTCAGCCTTCTGCTACACAGCTACCTGCAATACACGCTCCTATT 651

QY 1045 CTTCACTGCTGAAGAGCGTGTGCTGCGGCCGCCGACGCGCCGACGCGCTGGGCGCC 1104
Db 652 CAACAGCCTGTGAAGATGCTGTGTAATTAACAGGCTACAGCTGCTTTAGCACCT 711

QY 1105 TGCACCCCGCTGGTGGCTGCGCCGACGACCTGCCACTGAGAGCGCTGGAGAGAG 1164
Db 712 ACACACCCCTCTTGGATACACAGGCAATTCAAACCTGTTCAACCCAGTCTTTCTGTAG 771

QY 1165 GAGGCGAGCGGAGCCACACCCGCTGGATGTGACTATGGCGGCTCCGAGCGCAGGTGC 1224
Db 772 GGAACCGCTTCAATGTGACTGTGATGCCACTGTGCTGGAAGCTCCAAACTATCAAGA 831

QY 1225 CCACCGCTTCCGTATCCAAAGCACTGTGCTGCGCCGACGTAAGCTGTGAGCAGTACAGCGTG 1284
Db 832 CCACCAACACCTACCCAAACATCTGCTGACCAAAACCCCATCTGTTCTCCATACAG 891

QY 1285 GACCTGGAGAGCCTGTGCACCACTGTGCAGCAGAGCTGCGAGGGGCGCACTGATAGAC 1344
Db 892 TCAATCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTGCCCAAGGAAGATGAGAGTGAA 951

QY 1345 GGGAGTGACAGAGCCACAAAGGTGCGAGGCGAGACARAGCTGGCAGACACAAAAGCAG 1404
Db 952 AAGAGTTA-----TGAAATGTTGATGGGGATAAAGAAAGAACAG 996

QY 1405 ATTACAGACCTCCCGTGCCTGTCCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
Db 997 ATTACAACTTCACTTATTACTCTTAGGAAAAACAAGAAAGATGAAGAGCGAAGTCT 1056

QY 1465 CCATCAAGAGTACTCCCTTATGCTTCAATTTCTTCATCGAGACACACATCGAGAT 1524
Db 1057 CGTATTCAAGTATTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAT 1116

QY 1525 GTCATCAAAACCTACACAGAGAGGTGACGCGGAGGTACAGCTGAGCAGAGAAATGCC 1584
Db 1117 GTACTCAATCTCAGCAGCGCTCTACATCTGTAATAACAATTAGAGATGAATGATG 1176

QY 1585 AAAGCTGGGCTCTGTGAGCCGAGCAGGAGATGAGGAGATCTCTTACAGAGAGG 1644
Db 1177 CGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAA 1236

QY 1645 TCTAATCAACCGCTGAAGAGGCGCAAGATGGACAGTCCATGTTTGTGAAATCAAG 1704
Db 1237 TCTAATTACATCCGCTTTAAAAAGGCTAAATGGCAAGTCTATGTTTGTGAAGATAAG 1296

QY 1705 ACTCTAGGCATCGGTGCTTTGGGGAAGTGCCTCGCTTGTAAAGCTGCACACTCACGCT 1764
Db 1297 ACACTAGGATAGGAGCATTTGTTGAAGTCTCTAGCAAGAAAGTAGATCTTAGGCT 1356

QY 1765 CTGTACGCCATGAAGACTCTCAGGAAGAGAGTGCCTGAACCGGAATCAAGTGGCCAT 1824
Db 1357 TTGTATGCAACAAAACCTCTCGAAAGAAAGATGTTCTTTCGAAATCAAGTGCCTCAT 1416

QY 1825 GTCAAGCTGAGAGGAGCATCTCGCTGAAGCAGACATGAGTGGTGTCAACTCTAC 1884
Db 1417 GTTAAAGCTGAGAGAGATCTCGCTGAAGCTGCAATGAATGGTGTGCTCTATAT 1476

QY 1885 TACTCTTCCAGGCAAGGACAGCTGTACTTTGTGAGTGCATACATCCAGCGGGAT 1944
Db 1477 TATTCATCCAGATAGGACAAATTTACTTTGTAAGGACTACATTCCTGGGGGTAT 1536

QY 1945 ATGATGAGCTGCTGATCAGGATGGAGTCTTCCCTGAGCAGCTGGCCGCTTCTACAT 2004
Db 1537 ATGATGAGCTTATTAATAGAAATGGCATCTTCCAGAAAGTCTGCGCAGATTTACATA 1596

QY 2005 GCAGAGTTGACCTGGCCATTCGAAAGTGTCCACAGATGGGCTTTATCCACGGGACATC 2064
Db 1597 GCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGGTTTTTATTCATAGATAT 1656

QY 2065 AAGCCTGACACATCTCATCGACCTGGATGCTATATTAAAGCTGACAGATTTGGCCTC 2124
Db 1657 AAACCTGATATATTTGATGATCGTATGCTATATTAATTTGACTGACTTTGGCCTC 1716

QY 2125 TGCATGGATTCAGGTGGACTCACAATTCCTCAAGTACTACAGAAAGGGAACACATGAGA 2184
Db 1717 TGCATGGCTTCAGATGGACACACCATCTAAGTACTATCAGAGTGGTGAACCATCCAG 1776

QY 2185 CAGGACAGATGGAGCCGCTGACCTCTGGGAGAGTGTTCCAACTGCTGCTGTGAGAC 2244
Db 1777 CAAGATAGCATGGATTTTCAAGTAAATGGGGGGATCCCTCAAGCTGTGCTGTGAGAC 1836

QY 2245 AGGTTTAAAGCCTTGAGCAGAGGCGGAGAGAGCAGCAGAGTGGCTGGCAGATTTCT 2304
Db 1837 AGACTGAAGCCATTAGCGGAGAGCTGACCCAGCAGCAGGATGCTAGCATTTCT 1896

QY 2305 CTTGTGGGACACCAAAATTACATCTCGGAGGTGCTTCTCCGAAAGGGTACACGAG 2364
Db 1897 TTGGTTGGGACTCCCAATTTATTCACCTGAAGTGTGCTACGAACAGGATACACAG 1956

QY 2365 CTCCTGACTGGTGGAGCGTGGTGTGATCTCTTTGAGATGCTGGTGGGCGCGCT 2424
Db 1957 TTGTGTGATGGTGGAGTGTGGTGTATTCTTTTGTAAATGTTGGTGGGACAACTCT 2016

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokkwo, S., Oguh, M., Okwuonu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savary, G.,
 Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I.,
 Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, K., Wang, S.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G. and Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 57393)
 Worley, K.C.

Direct Submission
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 57393)
 Worley, K.C.

Direct Submission
 Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 17, 2002 this sequence version replaced gi:17941559.

----- Genome Center -----
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GBUE
 Center clone name: CH230-401
 Summary Statistics
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap: version 0.990329
 Consensus quality: 38843 bases at least Q40
 Consensus quality: 42803 bases at least Q30
 Consensus quality: 45893 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 33 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.

1 1327: contig of 1327 bp in length
 * 1328 1427: gap of unknown length
 * 1428 2480: contig of 1053 bp in length
 * 2481 2580: gap of unknown length
 * 2581 3874: contig of 1294 bp in length
 * 3875 3974: gap of unknown length
 * 3975 6236: contig of 2262 bp in length
 * 6237 6336: gap of unknown length
 * 6337 7588: contig of 1252 bp in length
 * 7589 7688: gap of unknown length
 * 7689 9085: contig of 1397 bp in length
 * 9086 9185: gap of unknown length
 * 9186 10268: contig of 1083 bp in length
 * 10269 10368: gap of unknown length
 * 10369 11925: contig of 1557 bp in length
 * 11926 12025: gap of unknown length
 * 12026 13508: contig of 1483 bp in length
 * 13509 13608: gap of unknown length
 * 13609 15071: contig of 1463 bp in length

RESULT 14	
E24614	
LOCUS	E24614
DEFINITION	1374 bp DNA linear PAT 18-JUN-2001
	warts Protein, polynucleotide encoding the protein, antisense
	polynucleotide thereof and antibody recognizing the protein.
E24614	
ACCESSION	E24614.1 GI:13024647
VERSION	JP 1999089580-A/2.
KEYWORDS	unidentified.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1374)
AUTHORS	Toshihiko,K., Shinichiro,T., Kyoko,N., Yasuyuki,N. and Hideyuki,S.
TITLE	warts Protein, polynucleotide encoding the protein, antisense
	polynucleotide thereof and antibody recognizing the protein

/organism="Drosophila melanogaster"		
/db_xref="taxon:7227"		
/map="100A2-5"		
gene	1..5360	
mRNA	/gene="warts"	
exon	1..5360	
exon	1..336	
exon	/gene="warts"	
exon	/number=1	
exon	337..1259	
exon	/gene="warts"	
CDS	/number=2	
	1137..4436	
	/gene="warts"	
	/function="required for the control of cell shape and proliferation"	
	/note="homolog of human myotonic dystrophy kinase"	
	/codon_start=1	
	/product="tumor suppressor"	
	/protein_id="AAA73959.1"	
	/db_xref="GI:755008"	
/translation="MHPAGEKRGGRPNKDYTAEALESIKODLIRFEVONHNNQNT PLRTATNGRNDALPDYHAKQPMPEPPSPASPADVPIPPPAIVGPGAGSLVSG VGVVGVANGRPKMTALPNKLIKPSIERDTASSHYLRCPALDAGAGSRSDS PHSHHTQPSRSRTVGNPGNGGSPSPSEVAPPAPPRNPRTASSAAATPPPPVPT SOAYKRRSPALNRPPIAIPPTORGNSPVTONGLNPOOQLTOQLSLNLYPGGGS GAVVEPPPYLLOGGAGGAAPPPPPSYATSMOSROSPQSOOSDRKSPSSGIYSAT SAGSPITVSLPAPLAKPQPVYQARSQPIINQSVKSTOVOKPVLQAVAFQPS SASNSPVLHAAAPSPYIPKSAVVOQQAHAHQHQQHQQHQQHQQHQQHQQHQQH LNSPNCLEPPSTAKSMQAAATVVOQQAHAHQHQQHQQHQQHQQHQQHQQHQQH LOAQERQERQERQERQERQERQERQERQERQERQERQERQERQERQERQERQER SNSNLATPPIPPAKYNNNSNTGANSNGSGSTGTASSSTCKKIKHASPPIPERK KISKEEERKEPIROYSPQAFKFMEOHENVIKSYRORTYKNOLEKEMHKGVL QOITKMLKNOHVAERDILAEADNNVWVVKLYSFQDKNLYFVMDIYIPGG AMTLKADVLKNOHVAERDILAEADNNVWVVKLYSFQDKNLYFVMDIYIPGG LMSLLIKLFEELARFYIAEVTCAVDSVHKMGFIHRDKPDLNLDIRGHIKLTDF GLCTFWNTIANSKYQENGRSDSPWEESENGPKPTVLERRRMRDRHVLHNS LVGTPNTIANSKYQENGRSDSPWEESENGPKPTVLERRRMRDRHVLHNS TLHPPOAELSREATDLIRLCASADRLGKSVDEKSHDFKGIIDFADMRKOKAPYI PEIKHPDTDSNEDVPDEKLRSNDSTSSGDDVDNDRTFPHGFETFRFFDDKQPP DMTDDQAPYY"		
exon	1260..3522	
	/gene="warts"	
exon	/number=3	
	3523..3643	
	/gene="warts"	
exon	/number=4	
	3644..3759	
	/gene="warts"	
exon	3760..3838	
	/gene="warts"	
exon	/number=6	
	3839..4034	
	/gene="warts"	
exon	/number=7	
	4035..5360	
	/gene="warts"	
	/number=8	
BASE COUNT	1562 a 1445 c 1385 g 968 t	
ORIGIN		
Query Match		17 98; Score 566.2; DB 3; Length 5360;
Best Local Similarity		64.48; Pred. No. 3.5e-115;
Matches 980; Conservative		0; Mismatches 493; Indels 48; Gaps 7;
Qy	1329 GGGCACTGATGACGCGAGTGCACAGAGCCAAAGGTGCGAGGAGGAGCAAGCTGG 1388	
Db	1329 GGGCGGAATAGCTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1388	
Qy	1389 CAGAGCAAAAAGCAGATTACAGCTCCCGGTCCTGTCGCGAAGATAGC----- 1440	
Db	1389 CAGAGCAAAAAGCAGATTACAGCTCCCGGTCCTGTCGCGAAGATAGC----- 1440	
Qy	1441 -AGAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTTACCAATT 1499	
Db	1441 -AGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTTACCAATT 1499	
Qy	1500 CTTTATGAGCAACACGCTGGAGAAATGTCTCAAAACCTTACCAGCAGAGGTGAG 1559	
Db	1500 CTTTATGAGCAACACGCTGGAGAAATGTCTCAAAACCTTACCAGCAGAGGTGAG 1559	
Qy	1560 GCTACAGCTGAGCAGAGAAATGGCCAAAGCTCTGTGAGGCGGAGGAGGAGCAGAT 1619	
Db	1560 GCTACAGCTGAGCAGAGAAATGGCCAAAGCTCTGTGAGGCGGAGGAGGAGCAGAT 1619	
Qy	1620 GAGGAAGATCTCTACCAAGAGGAGTCTAAGTACAAACCGCTGAAGAGGCGCAAGATGA 1679	
Db	1620 GAGGAAGATCTCTACCAAGAGGAGTCTAAGTACAAACCGCTGAAGAGGCGCAAGATGA 1679	
Qy	1680 CAAGTCCATGTTTGTGAAAATCAAGACTCTAGGCACTCGTGGGAGAGTGTGCTT 1739	
Db	1680 CAAGTCCATGTTTGTGAAAATCAAGACTCTAGGCACTCGTGGGAGAGTGTGCTT 1739	
Qy	1740 CGCTTGTAAAGCTGGA---CACTACGCTCTGTACCCATGAAGACTCTCAGGAAGAAGA 1796	
Db	1740 CGCTTGTAAAGCTGGA---CACTACGCTCTGTACCCATGAAGACTCTCAGGAAGAAGA 1796	
Qy	1797 TGTCTGTAACCGGAATCAAGTGGCCCATGTCAAGGCTCAGAGGAGCATCTCGGTGAAGC 1856	
Db	1797 TGTCTGTAACCGGAATCAAGTGGCCCATGTCAAGGCTCAGAGGAGCATCTCGGTGAAGC 1856	
Qy	1857 AGACAATGAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGAGCAGCTGTACTT 1916	
Db	1857 AGACAATGAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGAGCAGCTGTACTT 1916	
Qy	1917 TGTGATGACTACATACAGCGGGGATATGATGAGCTGTGATGAGTGGAGGAGTCTT 1976	
Db	1917 TGTGATGACTACATACAGCGGGGATATGATGAGCTGTGATGAGTGGAGGAGTCTT 1976	
Qy	1977 CCCTGAGCACCTGGCCCGCTTCTACATTCAGAGTTGACCTGGCCATTTGAAGTGTCCA 2036	
Db	1977 CCCTGAGCACCTGGCCCGCTTCTACATTCAGAGTTGACCTGGCCATTTGAAGTGTCCA 2036	
Qy	2037 CAAGATGGCTTTATCCACCGGAGCATCAAGCTTGACAACTCATCTCGACCTGGATGG 2096	
Db	2037 CAAGATGGCTTTATCCACCGGAGCATCAAGCTTGACAACTCATCTCGACCTGGATGG 2096	
Qy	2097 TCATATTAGCTGACAGATTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2156	
Db	2097 TCATATTAGCTGACAGATTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2156	
Qy	2157 GTACTTACC---AGAAAGGGAACACATGAGACAGCAGCAGCAGCAGCAGCAGCAGCAG 2213	
Db	2157 GTACTTACC---AGAAAGGGAACACATGAGACAGCAGCAGCAGCAGCAGCAGCAGCAG 2213	
Qy	2214 GGAGATGTTTCCAACTGCTGTGGAGACAGGTTAAAGACCCCTTGGAGCAGAGGCGCA 2273	
Db	2214 GGAGATGTTTCCAACTGCTGTGGAGACAGGTTAAAGACCCCTTGGAGCAGAGGCGCA 2273	
Qy	2274 GAAGCAGCAGCAGAGGTCCTTGGCAGATTTCTTTCGGGACACCAATATCATCGCTCC 2333	
Db	2274 GAAGCAGCAGCAGAGGTCCTTGGCAGATTTCTTTCGGGACACCAATATCATCGCTCC 2333	
Qy	2334 GGAGTGTCTTCCGCAAGGAGTACACGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2393	
Db	2334 GGAGTGTCTTCCGCAAGGAGTACACGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2393	
Qy	2394 TCTCTTTGAGATGCTTGGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2453	
Db	2394 TCTCTTTGAGATGCTTGGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2453	
Qy	2454 GCTGAAGTGTATCACTAGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2513	
Db	2454 GCTGAAGTGTATCACTAGGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2513	
Qy	4028 ACAAAGGTCATCACTAGGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4087	
Db	4028 ACAAAGGTCATCACTAGGAGAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4087	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 402.423 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagccaccgcgaagt.....aagagcactattttggggg 3155

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3155	100.0	3155	17 AAT42120	M-lats2 gene encod
2	3155	100.0	3155	21 AAZ51507	Mouse Lats2 (large
3	1749.6	55.5	5276	20 AAX87397	Human WART2 cDNA.
4	1749.6	55.5	5486	21 AAX59129	DNA encoding a tum
5	1748	55.4	5486	21 AAX59130	DNA encoding a tum
6	1717.2	54.4	3533	22 AAT17165	Human cancer-inhib
7	1293	41.0	1961	21 AAZ61160	CDNA SSI771A encod
8	1148.4	36.4	1498	21 AAZ61158	CDNA SSI771 encodi
9	1118.6	35.5	1912	23 ABK43457	DNA encoding novel

10	1035.2	32.8	2043	22 AAD03989	Human protein tyro
11	812.6	25.8	3213	21 AAZ51506	Mouse Lats (large
12	809.4	25.7	3213	17 AAT42119	M-lats gene encodi
13	760	24.1	3984	17 AAT42118	H-lats gene encodi
14	760	24.1	3984	21 AAZ51505	Human Lats (large
15	760	24.1	7382	20 AAX87396	Human WART1 cDNA.
16	758.8	24.1	1501	22 AAF75341	Human TGF-beta rec
17	747.8	23.7	2442	20 AAX32981	Human warts protei
18	676.6	18.4	1374	20 AAX32982	Human warts gene f
19	567.8	18.0	3319	23 ABL03169	Drosophila melanog
20	564.6	17.9	5720	17 AAT42117	Lats gene encoding
21	564.6	17.9	5720	21 AAZ51508	Drosophila melanog
22	551.8	17.5	1357	22 ABA08740	Human large tumour
23	546.6	17.3	676	22 AAS27179	CDNA encoding nove
24	546.6	17.3	676	23 ABK43776	DNA encoding novel
25	333	10.6	638	21 AAZ79942	Human colon cancer
26	298.8	8.2	2160	22 AAI66703	A. gossypii AG007
27	248.2	7.9	11187	23 ABL03168	Drosophila melanog
28	241.8	7.7	2255	23 ABL29695	Drosophila melanog
29	237	7.5	4983	22 AAF44629	Novel protein kina
30	235	7.4	1710	21 AAC43403	Arabidopsis thalia
31	235	7.4	2003	21 AAC49841	Arabidopsis thalia
32	233.8	7.4	1935	20 AAX06834	Disease associated
33	231.6	7.3	1894	21 AAC39567	Arabidopsis thalia
34	230	7.3	1922	21 AAC49426	Arabidopsis thalia
35	225.2	7.1	3018	17 AAT31452	Human Ndr serine/t
36	225.2	7.1	3018	24 ABK84773	Human cDNA differe
37	225.2	7.1	3800	22 ABA08668	Human NDR homologu
38	222.6	7.1	2101	17 AAT31451	Drosophila Ndr ser
39	220.2	7.0	2001	21 AAC36174	Arabidopsis thalia
40	214.4	6.8	678	22 AAZ79982	Human colon cancer
41	192.2	6.1	734	22 AAS27161	CDNA encoding nove
42	192.2	6.1	734	22 AAK56728	Human immune/haema
43	192.2	6.1	734	23 ABK43709	DNA encoding novel
44	192.2	6.1	734	23 ABK43985	DNA encoding novel
45	166.4	5.3	1788	19 AAV71036	mPKAc-green floure

ALIGNMENTS

RESULT 1
AAT42120
ID AAT42120 standard; cDNA: 3155 BP.

XX AAT42120;

AC AAT42120;

XX 31-JAN-1997 (first entry)

DT M-lats2 gene encoding large tumour suppressor.

XX Mouse; m-lats2 gene; large tumour suppressor; fetal brain;
protein-serine/threonine-kinase; cell proliferation; antisense;
dominant-negative; cancer; degenerative disorder; trauma;
growth deficiency; therapy; antitumour; vulnery; diagnostic;
transgenic plant; transgenic animal; growth; senescence; ds.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 1..2943
FT FT /*tag= a
FT FT /product= m-lats2 protein

XX WO9630402-A1.

PN Mouse Lats2 (large

PD Human WART2 cDNA.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-0411111.

XX (UYVA) UNIV YALE.

Qy	2821	TTCTTCGATGACAAACGGCTATCCCTTCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAGT	2880
Db	2821	TTCTTCGATGACAAACGGCTATCCCTTCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAGT	2880
Qy	2881	GCAGACCCAGGGGATCGCGACTTGGAAAGTGGCGCCGAGGGCTGCCAGCCGGTGTACGTG	2940
Db	2881	GCAGACCCAGGGGATCGCGACTTGGAAAGTGGCGCCGAGGGCTGCCAGCCGGTGTACGTG	2940
Qy	2941	TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATTCAGATTCCTTTTCAGAGACAA	3000
Db	2941	TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATTCAGATTCCTTTTCAGAGACAA	3000
Qy	3001	ACTCAAGCTTAGGAATCCTTCATTTTATTTAGTCTTGTAATAGGCAACAGGAAGAGTCAAC	3060
Db	3001	ACTCAAGCTTAGGAATCCTTCATTTTATTTAGTCTTGTAATAGGCAACAGGAAGAGTCAAC	3060
Qy	3061	ATGATTTCAAATTAGCCCTCTCAGGACCTTCACTGCATTAARACAGTATTTTTTAAAAA	3120
Db	3061	ATGATTTCAAATTAGCCCTCTCAGGACCTTCACTGCATTAARACAGTATTTTTTAAAAA	3120
Qy	3121	TTAGTACAGTATGAAAGACACTTATTTTGGGG	3155
Db	3121	TTAGTACAGTATGAAAGACACTTATTTTGGGG	3155

AA	AAZ51507;	
AC		
XX		
DT	21-JUN-2000	(first entry)

KW Mouse; Lats2; large tumour suppressor; cytostatic; vulnery; cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase; treatment; prevention; screening; cancer; skin; ovarian tumour; KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia; KW Luteinizing hormone hypogonadotropic hypogonadism; metaplasia; KW dysplasia; degenerative disorder; growth deficiency; physical trauma; KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX	Key	Location/Qualifiers
FH	Key	

FT /-lag= D
FT /note= "This region is erroneously repeated in the
FT mouse Lats2 DNA-sequence shown in figure 14"
FT

AA PN WO200010602-A1

18-AUG-1999; 99WO-US19068.

18 AUG 1958, 3003 0030337:
XX
XX

PI Xu T, Tao W, St John MAR., Fei X, Fukumoto RK, Zhang S;
PI Turenchalk GS, Stewart RA;

WPI; 2000-246496/21.
P-PSDB: AAY70392.

use of lats proteins, complexes of lats and cdc2 for treating cancer that is refractory to treatment by standard chemotherapy and radiation therapy and disorders associated with aberrant levels of cdc2 activity

Claim 44: pages 112-117; 134pp; English.

The present sequence is a DNA encoding mouse *Lats2* (large tumour suppressor) protein which is a cell overproliferation inhibitor and a negative regulator of cell cycle-dependent kinase *cdc2/cyclin A*. The present sequence is useful for treating cancer that is refractory to standard chemotherapy or radiation therapy such as hyperplasia, metaplasia, or dysplasia, and disorders associated with aberrant levels of *cdc2* activity. Conditions treated by promoting *cdc2* function include degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds. An animal model, preferably a mouse, in which a *lats* gene has been disrupted by homologous recombination, e.g. a *lats* knock-out mouse, is used for screening compounds that can be used to treat or prevent cancer, particularly skin cancer, soft tissue sarcomas and ovarian tumours, and disorders associated with pituitary dysfunction e.g. luteinizing hormone (LH) deficiency. The *lats* DNA is also used in gene therapy.

sequence 1155 bp. 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match	100.0%	Score 3155;	DB 21;	Length 3155;
Best Local Similarity	100.0%;	pred. No. 0;		
Matches 3155;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCCGATATCC	60	
Db	1	ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCCGATATCC	60	
Qy	61	CTCTCGCTTTTCCCAACGAGTCAGSCATTCGCGCAGCTGCAGAGGTGAACCGCGAGATG	120	
Db	61	CTCTCGCTTTTCCCAACGAGTCAGSCATTCGCGCAGCTGCAGAGGTGAACCGCGAGATG	120	
Qy	121	CTTCAGGAGTTGGTGAATCGCGCATGTGACAGAGATGGCTGGCAGAGCGCTCACGCG	180	
Db	121	CTTCAGGAGTTGGTGAATCGCGCATGTGACAGAGATGGCTGGCAGAGCGCTCACGCG	180	
Qy	181	ACGGCAGTAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAGATGGGTACTCTGGAC	240	
Db	181	ACGGCAGTAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAGATGGGTACTCTGGAC	240	
Qy	241	CCCAGGAATGAGCAGATTTGCGAGTCATCAGCAGACCTCCCCAGGAAAGGGCTGGCG	300	
Db	241	CCCAGGAATGAGCAGATTTGCGAGTCATCAGCAGACCTCCCCAGGAAAGGGCTGGCG	300	
Qy	301	TCCACCCCGGTGACTCGCGCGCCAGTTTCAGGGGCACAGGGGAAGCACTCCCATCTCTAC	360	
Db	301	TCCACCCCGGTGACTCGCGCGCCAGTTTCAGGGGCACAGGGGAAGCACTCCCATCTCTAC	360	
Qy	361	CACCACTGGGTGGTGCAAACTACGAGGGCCCGCCGCACTCGAGAGATCCGCGCAA	420	
Db	361	CACCACTGGGTGGTGCAAACTACGAGGGCCCGCCGCACTCGAGAGATCCGCGCAA	420	
Qy	421	TATTTAGACTTTCTCTTCCCTGGAGCGGAGCCGACCCAGCGTGCCAGGCTACACAG	480	
Db	421	TATTTAGACTTTCTCTTCCCTGGAGCGGAGCCGACCCAGCGTGCCAGGCTACACAG	480	
Qy	481	CATCTCTCCCAAGGGTACAGCAGCAGTAGAGCCAGTGGCACTTTTCCGGGCACACAC	540	
Db	481	CATCTCTCCCAAGGGTACAGCAGCAGTAGAGCCAGTGGCACTTTTCCGGGCACACAC	540	
Qy	541	TATGGTCGTGGTCATCTACTATCGGAGCAGTCTGGGTATGGGTGTCAGCGCAGTTCCTCC	600	
Db	541	TATGGTCGTGGTCATCTACTATCGGAGCAGTCTGGGTATGGGTGTCAGCGCAGTTCCTCC	600	
Qy	601	JTCCAGAACACAGACGCCACACAGATGCCTATTTCACAGATGCCAAGGCCACCGGTCG	660	

Db 1681 AAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCCCTTTGGGAAAGTGTCCCTC 1740
 QY 1741 GCTTGTAAAGCTGGACACTCAGCTCTGTAGCCATCAAGACTCTCAGGAAGAGGATGTG 1800
 Db 1741 GCTTGTAAAGCTGGACACTCAGCTCTGTAGCCATCAAGACTCTCAGGAAGAGGATGTG 1800
 QY 1801 CTGAACCGGAATCAAGTGGCCCATGTCAGGCTGAGAGGACATCTGCTGAAGCAGAC 1860
 Db 1801 CTGAACCGGAATCAAGTGGCCCATGTCAGGCTGAGAGGACATCTGCTGAAGCAGAC 1860
 QY 1861 AATGAGTGGGTGCTAACTCTACTACTCTCTTCCAGGACAAAGACAGCTGTACTTTGTG 1920
 Db 1861 AATGAGTGGGTGCTAACTCTACTACTCTCTTCCAGGACAAAGACAGCTGTACTTTGTG 1920
 QY 1921 ATGGACTACATACACAGCGGGGATATGATGAGCCTCTGATCAGGATGAGGTCTTCCCTC 1980
 Db 1921 ATGGACTACATACACAGCGGGGATATGATGAGCCTCTGATCAGGATGAGGTCTTCCCTC 1980
 QY 1981 GAGCACTGGCCCGCTTCTACATTCAGAGTGTGACCTTGGCCATTTGAAAGTGTCCACAAG 2040
 Db 1981 GAGCACTGGCCCGCTTCTACATTCAGAGTGTGACCTTGGCCATTTGAAAGTGTCCACAAG 2040
 QY 2041 ATGGGCTTTATCCACGGGACATCAAGCTGACAACTACTATCAGCCTGGATGCTCAT 2100
 Db 2041 ATGGGCTTTATCCACGGGACATCAAGCTGACAACTACTATCAGCCTGGATGCTCAT 2100
 QY 2101 ATTAAGCTCACAGATTTTGGCCTCTGCACTGGATTCAGTGGACTCACAAATCCCAAGTAC 2160
 Db 2101 ATTAAGCTCACAGATTTTGGCCTCTGCACTGGATTCAGTGGACTCACAAATCCCAAGTAC 2160
 QY 2161 TACCAGAAAGGAACACATGACAGGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
 Db 2161 TACCAGAAAGGAACACATGACAGGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
 QY 2221 GTTTCACACTGCTGTGGAGACAGGTTAAAGACCTCGAGCAGAGGCGCAGAGCAG 2280
 Db 2221 GTTTCACACTGCTGTGGAGACAGGTTAAAGACCTCGAGCAGAGGCGCAGAGCAG 2280
 QY 2281 CACCAGAGTGTGCTGACATCTCTGTCGGGACACAAATACATCCTCCGGAGGTG 2340
 Db 2281 CACCAGAGTGTGCTGACATCTCTGTCGGGACACAAATACATCCTCCGGAGGTG 2340
 QY 2341 CTTCTCCGCAAAAGGTTACACGAGCTCTGTACTGTGGAGGCTGGGTGTGATCTCTTT 2400
 Db 2341 CTTCTCCGCAAAAGGTTACACGAGCTCTGTACTGTGGAGGCTGGGTGTGATCTCTTT 2400
 QY 2401 GAGATGCTGGTGGGACGCGCTTCTTGGCCCGCCACCCACAGACGAGCAGCTGAAG 2460
 Db 2401 GAGATGCTGGTGGGACGCGCTTCTTGGCCCGCCACCCACAGACGAGCAGCTGAAG 2460
 QY 2461 GTGATCACTGGGAGACGCTGTGATATCCCTACGAGTGGGCTCAGCGCTGAGGCC 2520
 Db 2461 GTGATCACTGGGAGACGCTGTGATATCCCTACGAGTGGGCTCAGCGCTGAGGCC 2520
 QY 2521 CGAGACCTCATCAGGAAGCTGTGCTGCGCGGTGACTGCGGCTGGGAGGATGGGCA 2580
 Db 2521 CGAGACCTCATCAGGAAGCTGTGCTGCGCGGTGACTGCGGCTGGGAGGATGGGCA 2580
 QY 2581 GATGACCTCAAGGCACACCCGTTCTCAACACCATCGACTTTCCCGTGACATCCGAAG 2640
 Db 2581 GATGACCTCAAGGCACACCCGTTCTCAACACCATCGACTTTCCCGTGACATCCGAAG 2640
 QY 2641 CAGGTGACCTACCTGCTCCACCATCAGCCACCCATGACACCTCCCAATTTGACCG 2700
 Db 2641 CAGGTGACCTACCTGCTCCACCATCAGCCACCCATGACACCTCCCAATTTGACCG 2700
 QY 2701 GTGGATGAAGAAAGCCCTTGGCAGCAGGCGGAGGAGAGGCGCCAAAGCCTGGGACAG 2760
 Db 2701 GTGGATGAAGAAAGCCCTTGGCAGCAGGCGGAGGAGAGGCGCCAAAGCCTGGGACAG 2760
 QY 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGACAGCGCTTCTATGAGTTCACCTTCCGACG 2820
 Db 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGACAGCGCTTCTATGAGTTCACCTTCCGACG 2820

QY 2821 TTCTTCGATGACAAAGGCTATCCCTTCGGTGCCTGAGCCCTCAGAGCCCGCAGAGAT 2880
 Db 2821 TTCTTCGATGACAAAGGCTATCCCTTCGGTGCCTGAGCCCTCAGAGCCCGCAGAGAT 2880
 QY 2881 GCAGACCCAGGGATGCGGACTTGGAGGTGCGGCCGAGGCTGCCAGCGGTGTACGTG 2940
 Db 2881 GCAGACCCAGGGATGCGGACTTGGAGGTGCGGCCGAGGCTGCCAGCGGTGTACGTG 2940
 QY 2941 TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATGAGATTTCTTTTCAGAAGACAA 3000
 Db 2941 TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATGAGATTTCTTTTCAGAAGACAA 3000
 QY 3001 ACTCAAGCTTAGGAATCCCTTCATTTTCTGTGTAATGGCAACAGAGAGTCAAC 3060
 Db 3001 ACTCAAGCTTAGGAATCCCTTCATTTTCTGTGTAATGGCAACAGAGAGTCAAC 3060
 QY 3061 ATGATTTTCAAAATAGCCCTCTGAGGACCTTCACTGCATTTAAACAGTATTTTAAAAAA 3120
 Db 3061 ATGATTTTCAAAATAGCCCTCTGAGGACCTTCACTGCATTTAAACAGTATTTTAAAAAA 3120
 QY 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155
 Db 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155

RESULT 3

AA87397
 ID AA87397 standard; cdna; 5276 BP.

XX AC AA87397;
 XX DT 08-OCT-1999 (first entry)
 XX DE Human WART2 cdna.

KW WART2; hWART2; WART orthologue; human; signal transduction;
 protein kinase; cancer; tumour; diagnosis; therapy; ss.

XX OS Homo sapiens.

Key	Location/Qualifiers
5'UTR	1..374
FT	/*tag= a
FT	375..3641
FT	/*tag= b
FT	3642..5276
FT	/*tag= c

XX PN W09937787-A2.

XX PD 29-JUL-1999.

XX PF 20-JAN-1999; 99WO-US01145.

XX PR 21-JAN-1998; 98US-0072023.

XX PA (SUGEN-) SUGEN INC.

XX PI Flanagan P, Plowman GD;

XX DR WPI; 1999-458698/38.

XX PT P-PSDB; AA06527.

PT New nucleic acid encoding human orthologs of Drosophila WART
 proteins, used to identify specific modulators for treating cancer
 or for diagnosis

XX PS Claim 1; Page 122-123; 137pp; English.

XX CC This is the nucleotide sequence of a cDNA clone coding for a human
 orthologue, i.e. hWART2 (see AA06527), of Drosophila non-receptor
 serine/threonine kinase WART. hWART2 cDNA was isolated from a

QY	733	CTGGGGCCCA	CCCAATTACAT	TGTGTGGCCAC	CCCGGGTCC	CCACGT	TTTACT	TGCGAAAGC	792	
Db	1341	CCCGGCCCA	CCCAAGTCAT	TGTGCTGGGCT	CCCGAC	CCAGGT	TGTCG	CCAGCGACG	1400	
QY	793	TCTGCACAG	CGTGTGCTGG	CACCTTCCAG	GAACAGCCT	CAATGT	GACTTGT	TACGAGCTG	852	
Db	1401	CCCGGAGAC	CGTGTCTACT	CCCTTCGG	GAACACGCT	CAACGT	TGACCT	TGTATGAATTG	1460	
QY	853	GG-----	CTCACGGT	GCCTTGCT	CTGCAGCT	CCACT	TGCGAC	GC	CGGACTCGCTGCAG	906
Db	1461	GGCAGCA	CCCTCCGTC	CCAGCAGT	GC	CGCTGCC	CACT	TGCGGACT	CCCTGCGAG	1520
QY	907	AAGCAGG	CTAGAAC	CGCTCGCGC	---CC	ATGTG	CGCTTTT	TCGGGCTG	954	
Db	1521	AAGCCGG	CGCTTGAG	GGCGCGCG	CGCGCG	ACGTG	GCCTT	TCGGGCT	GACTGCC	1580
QY	955	CCGACG	AGGACCA	CTCTTTCA	CAACCCACA	AC-----	-----	-----	CT	990
Db	1581	CCGACG	AGGACCA	CTCTTCA	CAGCCAC	CCCGGG	CGGCT	CGGCT	GGAAGGC	1640
QY	991	GAGCC	CTACTGCC	CGGCCCA	CACGGT	CA	CGCGT	GACGG	CCGACACAT	1050
Db	1641	GAGCC	CTCCCTGCC	CGCCCA	CACGGT	GTCA	CGCG	CGCACAT	CTTGCAC	1700
QY	1051	CTTGT	GAAGAG	CGTGTG	CTGCG	CCGAG	CCCGAC	AGCCTG	GGGGCCCTCG	1110
Db	1701	CGSGT	GAAGAG	CGTGTG	CTGAG	CCCGAG	CCCGAC	AGCCTG	GGGGCCCTCG	1760
QY	1111	CCGCG	CTGGT	GTGCTG	CGC-----	-----	CCAC	AGCAC	CTGCCACTG	1149
Db	1761	CCGCG	CTGGTGT	CGCGG	CGCTGCC	CGCGCG	CGCGCG	CGCGCG	CGCGCTCG	1820
QY	1150	AGCT	TGGAC	CAAGA	GA-----	-----	GGC	GAGG	CGCCAC	1191
Db	1821	GCTT	TGGAC	CAAG	GAGG	AGATG	CCCTG	GGCG	CGGAGG	1880
QY	1192	GATG	TGGACT	ATGG	CGGCTTC	GAGCG	CAAGT	GTGCC	ACCGCTT	1251
Db	1881	GACGT	TGGAGT	ACG	GAGG	CCAGC	CGGAG	GTCC	CGCTTCA	1940
QY	1252	CTGT	CTCCCA	GTAA	GTCTG	GACAGT	TACAG	CGTGG	ACCTGG	1311
Db	1941	CTGT	CTGCCA	CAAGT	TCG	GACAG	CTAC-----	GACCT	TGGAC	1994
QY	1312	CAGC	AGACT	TCG	GAGGGG	---CA	CTGAT	CTAG	ACGGG	1368
Db	1995	GAGC	AGAC	CTCCG	TGCGGG	CCCCA	AGC	CGCG	CGCGAC	2054
QY	1369	GC	GAAGG	GAGACA	AGCTG	GC	GAGAC	AAAG	CAGATT	1428
Db	2055	GCC	AAAGG	GAGACA	AGG	GA	GAATTA	AAAG	CAGATT	2114
QY	1429	CGC	AAAGAT	AGC	AGAGAT	GAAG	AGAG	AGTCT	CGCAT	1488
Db	2115	CGC	AAAAAC	AGC	AGAGAC	GAAG	AGAG	AGTCT	CGCAT	2174
QY	1489	GCCTT	CAAA	TTCTAT	CGG	AGAC	AAAG	CAGATT	TAGAC	1548
Db	2175	GCCTT	TAAGT	TTCTAT	CGG	AGAC	AGC	AGTGT	GGAAT	2234
QY	1549	GT	CAG	CGG	AGGCTT	AC	AGCTT	GAG	CAG	1608
Db	2235	GTT	AAC	CGG	AGGCTT	AC	AGCTT	GAG	CAG	2294
QY	1609	CAG	AGC	AGAT	GAG	AGAT	CTCTT	AC	CAG	1668
Db	2295	CAG	AGC	AGAT	GCG	AAAGT	CTCTT	AC	CAG	2354
QY	1669	GCC	AAAT	GGAC	AAAGT	CCAT	GT	TTGT	GAAAA	1728
Db	2355	GCC	AAAT	GGAC	AAAGT	CCAT	GT	TTGT	GAAAA	2414
QY	1729	GA	AGT	GT	GCCT	GTG	TAA	GTG	GAC	1788

Db 2415 GAAGTGTGCTTGTGTAAGTGGACATCAAGCCCTGACGCCATGAAGACCCCTAAGG 2474
 QY 1789 AAGAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTCAGAGGACATCCCTG 1848
 Db 2475 AAAAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTCAGAGGACATCCCTG 2534
 QY 1849 GCTGAAGCAGACAATGAGTGGGTGCTCAAACTCTACTACTCTCTCAAGCAGACAGACAG 1908
 Db 2535 GCCGAGGACAGACAATGAGTGGGTGCTCAAACTCTACTACTCTCTCAAGCAGACAGACAG 2594
 QY 1909 CTGTACTTGTGATGACTACATACACAGCGGGGATATGATGAGCTGTCTGATCAGGATG 1968
 Db 2595 CTGTACTTGTGATGACTACATACACAGCGGGGATATGATGAGCTGTCTGATCCGGATG 2654
 QY 1969 GAGGTCTTCCCTGAGCAGCTGCGCGCTTCTACATTCAGAGTGTGACCTGCGCATTTGAA 2028
 Db 2655 GAGGTCTTCCCTGAGCAGCTGCGCGCTTCTACATTCAGAGTGTGACCTGCGCATTTGAG 2714
 QY 2029 AGTGTCCACAAGATGGCTTTATCCACCGGACATCAAGCCTGACAACTACTCTATCTGAC 2088
 Db 2715 AGTGTCCACAAGATGGCTTTATCCACCGGACATCAAGCCTGACAACTACTCTATCTGAC 2774
 QY 2089 CTGGATGGTCATTAAGCTGACAGATTTGGCCTCTGCACTGGATTCAGTGGGACTCAC 2148
 Db 2775 CTGGATGGTCATTAAGCTGACAGATTTGGCCTCTGCACTGGGTTCAAGTGGGACTCAC 2834
 QY 2149 AATTCCAAGTACTACACAGAGGACACATGACAGCAGACAGATGAGCCGCGGTGAC 2208
 Db 2835 AATTCCAAGTACTACACAGAGGACACATGACAGCAGACAGATGAGCCGCGGTGAC 2894
 QY 2209 CTCTGGGAGATGCTTTCAACTGCTGCTGGAGACAGGTTAAAGACCCCTGGAGCAGAGG 2268
 Db 2895 CTCTGGGAGATGCTTTCAACTGCTGCTGGGACAGGCTGAGACCCCTAGACAGAGG 2954
 QY 2269 GCGCAGACAGCAGCAGAGGCTCTGCGACATTTCTTTGTCGGGACACCAATATACATC 2328
 Db 2955 GCGCAGACAGCAGCAGAGGCTCTGCGACATTTCTTTGTCGGGACACCAATATACATC 3014
 QY 2329 GCTCCGGAGGCTCTCTCGCAAGAGGTACACGAGCTCTGTGACTGTGGAGGCTCGT 2388
 Db 3015 GCACCGAGGCTCTCTCGCAAGAGGTACACGAGCTCTGTGACTGTGGAGGCTCTGGA 3074
 QY 2389 GTGATTCCTTTGAGATGCTGTTGGCAGCGCCCTTTCTGGCCCCCACCACCCACAGAG 2448
 Db 3075 GTGATTCCTTTGAGATGCTGTTGGCAGCGCCCTTTTGGCACCTACTCCACAGAA 3134
 QY 2449 AGCAGCTGAAGGTGATCAACTGGGAGACGACGCTGCATATCCCTACGAGGTGAGGCTC 2508
 Db 3135 ACCCAGCTGAAGGTGATCAACTGGGAGACGACGCTGCATATCCAGCCCGAGGTGAAGCTG 3194
 QY 2509 AGCGCTGAGGCGGAGACCTCATACAGAAAGCTGTGCTGGCGGCTGACTGCGCGCTGGGC 2568
 Db 3195 AGCGCTGAGGCGGAGACCTCATACAGAAAGCTGTGCTGGCGGCTGACTGCGCGCTGGGC 3254
 QY 2569 AGGATGGGGAGATGACTCAAGCAGACACCGTCTTCTTCAACACCATGACTTTTCCCGT 2628
 Db 3255 CGGAATGGGGCGATGACTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3314
 QY 2629 GACATCCGAAAGAGGCTGACCTTACCTTCCCTCCCATCATGAGCCACCCCATGACACCTCC 2688
 Db 3315 GACATCCGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3374
 QY 2689 AATTTTGACCGGTGGATGAAGAAAGCCCTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2748
 Db 3375 AATTTTGACCGGTGGATGAAGAAAGCCCTTGGCAGGATGAGGAGGAGGAGGAGGAGGAG 3434
 QY 2749 GCTGGGACAGCTGGCTTCCCGCAGCAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2808
 Db 3435 GCTGGGACAGCTGGCTTCCCGCAGCAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3494
 QY 2809 ACCTTCGAGGCTTCTGATGACAAAGGCTATCCCTTCCGGTGCCTGAGGAGGAGGAGGAG 2866

Db 3495 ACCTTCCGAAGGTCTTTTGTGATGACAATGCTACCCCTTTTCGATGCCAAAGCCTTCAGGA 3554
 QY 2867 -----ACCCGCGAGAGTGCAGACCCAGGGGATCGGACTTGCAGAGTGGCGCC 2916
 Db 3555 GCAGAGCTTACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3614
 QY 2917 GAGGCTGCCAGCCGCTGTACGTGA 2942
 Db 3615 GAAGCTGCCAGCCTGTGTACGTGA 3640
 RESULT 4
 AAA59129
 ID AAA59129 standard; DNA; 5486 BP.
 XX
 AC AAA59129;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS1.
 KW Human: growth hormone inhibited tumour suppressor protein: hGHITS1;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 387..3653
 FT /*tag= a
 FT /product= "tumour suppressor protein"
 XX
 PN EP1022333-A1.
 XX
 PD 26-JUL-2000.
 XX
 XX 07-OCT-1999; 99EP-0119199.
 XX
 PR 25-JAN-1999; 99JP-0016223.
 XX
 PA (JCRP-) JCR PHARM CO LTD.
 XX
 PI Koga J, Kono K, Zolotar'ov FN;
 XX
 DR WPI: 2000-516013/47.
 DR P-PSDB; AAB07633.
 XX
 PT New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 PT diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 PT nephropathy or cardiopathy
 XX
 PS Claim 1; Page 18-28; 59pp; English.
 XX
 CC The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnostic pharmaceutical preparations can be used for examining
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.
 XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;
 Query Match 55.5%; Score 1749.6; DB 21; Length 5486;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 2389; Conservative 0; Mismatches 544; Indels 153; Gaps 16;
 QY 1 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATTC 60

Db 576 ATGAGAGCCACCCAAAGTTGCGACCTTATCAAAAGCCTTGAGGGAATACAGATTCC 635
QY 61 CTCCTGCTTTTCCCAACAGTTCAGGAGCTTCGGAGCTGCGAGAGTGAACCGCAGATG 120
Db 636 TTGTTGCTTTTGTCTAATCAATCGGCGACCTTCAGCTGCAAGTGAACCGCAATG 695
QY 121 CTTGAGGAGTTGGTGAATCGGCGATGTGACAGAGATGGCTGGCAGAGCGCTACGAG 180
Db 696 CTTGAGGAACTGGTGAACAGAGATGCGACAGGAGATGGCTGGCGGAGCTCAAGCAG 755
QY 181 ACGGCGAGTAGAGATCAAGAGTTCAGAGTTCATCAAGAGAGCTCCCAAGGAAAGGCTGCG 300
Db 756 ACTGGCAGCAGGAGATGAGCGCCCTCGGAGTCAATAGCAGACCTCCCAAGGAAAGGCTCATG 875
QY 241 CCCAGGAATGACAGATTGTGCGAGTTCATCAAGAGAGCTCCCAAGGAAAGGCTGCG 300
Db 816 CCGAGGAATGACAGATTGTGCGGCTCATTAAGCAGACCTCCCAAGGAAAGGCTCATG 875
QY 301 TCCACCCCGGTGACTCGGCGGCCAGTTTCAGAGGACAGAGGGAAGCACTCCCACTCTAC 360
Db 876 CCAACCCCGGTGACTCGGCGGCCAGTTTCAGAGGACAGAGGGAAGCACTCCCACTCTAC 360
QY 361 CACAGCTGGGTGTCGAACTACAG-----GGCCCGCGCGCA 399
Db 936 CACAGCTGAGCGGTACCGCTACAGGCGGCCAAAGTTCGCGCTGAGCGGCCCGCACGCG 995
QY 400 CTTGAGGAGATGCGCGCGCAATATTAGACTTCTTCCCTGAGCGCGGAGCGGCGAC 459
Db 996 CTTGAGGAGATGCGCGCGCTGACTGACTTCTTCCCGGAGTGGCGGCCCGCGG 1055
QY 460 CACGCTGCCAGGCTCACAGATCTCCCAAGGATACAGCAGCAGTAGAGCCAAAGT 519
Db 1056 CCGCGCCACCAAG-----CACAGCAGCCCAAGGCTACGCTGCGCAGCGTAGAGCAGCA 1112
QY 520 GCG-----CACTTTCGCGGACACACTATGCTGCTGCTACTACT-----ATCG 564
Db 1113 GGGCAGACTTCCCGCTCAGGCGCGGCTACGCGGCGCGCGCTGCTGCTGCTGGG 1172
QY 555 GAGCAGTCTGGGTATGGGTGACGCGAGTTCCTCTTCCAGAACAGAGCGCCACAGAT 624
Db 1173 GAACCCCTGGGCTACGAGTACGCGAGCGCCCTCTTCCAGAGCAGAGCGCGCGGAG 1232
QY 625 GCG-----TATTCAGCATGCGC---AAGCCCGAGGTGGCGCTCC---CGCCAGCCTC 672
Db 1233 ACCGGGGTTACGCGCGCTGCGCAGAGGCGCGAGGAGCGCGCGCGCGCTC 1292
QY 673 ACCTTTCCTGCGCATGCTGGGCTGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Db 1293 GCTTTCACCCCTGCGCGCGGCTCTACGTGCGCGCACCCACACAGCAGCGCGGCT 1352
QY 733 CTTGGGCGCCACCATTTACATGTTGGGACCGCGGCTCCCACTTTACTTGGCGAAAGC 792
Db 1353 CCGCGCGCCACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1412
QY 793 TCTCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
Db 1413 CCGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1472
QY 853 GG-----CTCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 906
Db 1473 GCGAGCCTTCCGTCAGCAGTGGCGGCTGCCACCTGCGCGCGCGGCGGCTCCCTGCGAG 1532
QY 907 ACGAGGCTTAGAGCTCGCGGCG-----CGCATGCTGCTGCTGCTGCTGCTGCTGCT 954
Db 1533 ACGCGGCGCTGAGGCG 1592
QY 955 CCGAGGAGCAACTCTTCAACACCCCAAC-----CT 990
Db 1593 CCGAGGAGCAACTCTTCAACACCCCAAC-----CT 1652
QY 991 GAGCCTCTGCTGCG 1050
Db 1050 GAGCCTCTGCTGCG

Db 1653 GAGCCTCTGCTGCG 1712
QY 1051 CTTGTAAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1110
Db 1713 CCGGTGAAGAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1772
QY 1111 CCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1149
Db 1773 CCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1832
QY 1150 AGCCTGGAGACCAAGG-----GGCAGCGCAGGCGCCACACCCGCTG 1191
Db 1833 GCTTGGAGCGCAAGGAGGAGCATGCCCTGGCTGGCGCGCGAGGCGCTTCCCGCTG 1892
QY 1192 GATGTGACTATGCGGCTCGCAGCGAGTGGCCCGCGCTGCTGCTGCTGCTGCTGCTGCT 1251
Db 1893 GACGTGAGTACGAGGCGCCAGACCGGAGTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1952
QY 1252 CTGCTGCCAGTAACTGCTGAGCAGTACAGCGTGGACCTGGACAGCCTGCTGCTGCTGCT 1311
Db 1953 CTGCTGCGCAAGTCTGAGCAGTAC-----GACCTGGACAGCCTGCTGCTGCTGCT 2006
QY 1312 CAGCAGAGTCTGCGAGGCGG---CACTGATCTAGCAGGAGTGCACAGAGCCACCAAGGT 1368
Db 2007 GAGCAGAGCTCTGCTGCGGCGCCCAACGAGCGCGGCGCGGAGAGCGCGCAAGC 2066
QY 1369 GCGAAGGAGACAAAGCTGCGCAGAGCAAAAGAGATTCAGACCTCCCGGCTGCTGCT 1428
Db 2067 GCGAAGGAGGAGCAAGCGCGGAAAGGATAAAGAGCAGATTTCAGACCTCTGCTGCT 2126
QY 1429 CGCAAGAAATAGCAGAGATGAAGAGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 1488
Db 2127 CGCAAAACAGCAGAGACGAGAGAGAGAGTACGCGATCAAGAGCTACTCGCCATAC 2186
QY 1489 GCCTTCAATCTTCATGAGAGCAACAGCTGGAGAAATGTATCAAAACCTTACCAAGAG 1548
Db 2187 GCCTTAAAGTCTTCTGAGAGCAGCAGCTGAGAAATGTATCAAAACCTTACCAAGAG 2246
QY 1549 GTGAGCGGAGGCTACAGCTGAGAGAGGAAATGCGCAAGCTGCTGCTGCTGCTGCTGCT 1608
Db 2247 GTTAAACGAGAGCTGAGCTGAGAGAGAAATGCGCAAGCTGCTGCTGCTGCTGCTGCT 2306
QY 1609 CAGGAGCAGATGAGAGAGATCTCTTACAGAGAGAGTCTAACTACAAACCGCTGAGAGG 1668
Db 2307 CAGGAGCAGATGCGGAAGATCTCTTACAGAGAGAGTCTAACTACAAAGCTTAAAGAG 2366
QY 1669 GCGAAGATGAGCAAGTCCATGTTGTAATCAAGAGTCTAGGATCGGCTGCTGCTGCTGCT 1728
Db 2367 GCGAAGATGAGCAAGTCTATGTTGTAAGATCAAAACCTGCGGATCGCTGCTGCTGCTGCT 2426
QY 1729 GAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1788
Db 2427 GAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2486
QY 1789 AGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1848
Db 2487 AAAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2546
QY 1849 GCTGAGAGCAGCAATGAGTGGTGGTCAACTCTACTACTCTTCTTCCAGGAGAGCAGCAG 1908
Db 2547 GCGAGGAGCAGCAATGAGTGGTGGTCAACTCTACTACTCTTCTTCCAGGAGAGCAGCAG 2606
QY 1909 CTGACTTGTGATGAGTACATACAGGCGGGGATATGATGAGCTGCTGCTGCTGCTGCTGCT 1968
Db 2607 CTGACTTGTGATGAGTACATCCCTGGTGGGAGATGATGAGCTGCTGCTGCTGCTGCTGCT 2666
QY 1969 GAGTCTTCCCTGAGCAGCTGCGCGCTTCTACATTCGAGTTCGAGCTGCTGCTGCTGCTGCT 2028
Db 2667 GAGTCTTCCCTGAGCAGCTGCGCGCTTCTACATTCGAGTTCGAGCTGCTGCTGCTGCTGCT 2726
QY 2029 AGTGTCCCAAGATGGCTTTTATCCAGCGGAGATCAAGCTGAGCAACATCTCTATCCAG 2088
Db 2727 AGTGTCCCAAGATGGCTTTTATCCAGCGGAGATCAAGCTGAGCAACATCTCTATCCAG 2786

QY 2089 CTGGATGTCATATTAAGCTGACAGATTTGGCTCTGCACTGGATTGAGTGGGACTCAC 2148
 Db 2787 CTGGATGTCATATTAAGCTGACAGATTTGGCTCTGCACTGGATTGAGTGGGACTCAC 2846
 QY 2149 AATTCAGAGTACTACAGAAAGGAGACCATGAGCAGGACAGATGAGCCGGTGC 2208
 Db 2847 AATTCAGAAATATACAGAAAGGAGGACCATGAGCAGGACAGATGAGCCGGTGC 2906
 QY 2209 CTCTGGGAGAGTGTTCCTCAACTGCTGCTGGAGACAGGTAAAGACCTGAGCAGAGG 2268
 Db 2907 CTCTGGGATGATGCTCTCAACTGCTGCTGGAGACAGGTAAAGACCTGAGCAGAGG 2966
 QY 2269 GCGCAGAAAGCAGCAGAGAGTGCCTGGCACAATCTTGTGGGACACCAAAATACATC 2328
 Db 2967 GCGCGGAAAGCAGCAGAGAGTGCCTGGCACAATCTTGTGGGACACCAAAATACATC 3026
 QY 2329 GCTCGGAGGCTCTCTCGGAAAGGTTACAGCAGCTCTGTGACTGTGGAGCGTCGGT 2388
 Db 3027 GCACCGAGGTGCTCTCTCGGAAAGGTTACAGCAGCTCTGTGACTGTGGAGCGTCGGT 3086
 QY 2389 GTGATCTCTTGTGAGATGCTGCTGGCAGCGGCTTCTTGGCCGCCACCCACAGAG 2448
 Db 3087 GTGATCTCTTGTGAGATGCTGCTGGCAGCGGCTTCTTGGCCGCCACCCACAGAG 3146
 QY 2449 AGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGTCATATCCCTACGAGGTGAGGCTC 2508
 Db 3147 ACCCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGTCATATCCCTACGAGGTGAGGCTC 3206
 QY 2509 AGCGTGAAGGCGGAGAGCTCTATCAGAGAGTGTGCTGGGCGGCTGACTGCGCGCTGGGC 2568
 Db 3207 AGCGTGAAGGCGGAGAGCTCTATCAGAGAGTGTGCTGGGCGGCTGACTGCGCGCTGGGC 3266
 QY 2569 AGGGATGGGAGAGTCACTCAAGCAGCAGCGCTTCTCAACACCATGACTTTTCCCGT 2628
 Db 3267 CGGAATGGGCGGAGTCACTCAAGCAGCAGCGCTTCTCAACACCATGACTTTTCCCGT 3326
 QY 2629 GACATCCGAAAGAGGCTGACCTTACGTGCCACCATCAGCAGCAGCTGACACCTCC 2688
 Db 3327 GACATCCGAAAGAGGCTGACCTTACGTGCCACCATCAGCAGCAGCTGACACCTCC 3386
 QY 2689 AATTTGACCGGCTGATGAAGAAGCCCTTGGCAGAGGCGGAGAGAGCGGAGAG 2748
 Db 3387 AATTTGACCGGCTGATGAAGAAGCCCTTGGCAGAGGCGGAGAGAGCGGAGAG 3446
 QY 2749 GCCTGGGACAGCTGGCTCCCGCAGCAGAGCATCAGACAGCGCTTCTATCAGTTC 2808
 Db 3447 GCCTGGGACAGCTGGCTCCCGCAGCAGAGCATCAGACAGCGCTTCTATCAGTTC 3506
 QY 2809 ACCTTCGAGGCTTCTTCGATGACAAAGGCTATCCCTTCCGGTGGCCGAGCCCTCAG 2866
 Db 3507 ACCTTCGAGGCTTCTTCGATGACAAAGGCTATCCCTTCCGGTGGCCGAGCCCTCAG 3566
 QY 2867 -----AGCCCGAGAGTGCACAGCCAGGAGTGGGATGGGACTTGAAGGTGCGGC 2916
 Db 3567 GCAGAGGCTTACAGGCTGAGAGCTCAGATTGAGAAAGCTCTGATCTGGTGGATCAGACT 3626
 QY 2917 GAGGCTGCCAGCGGCTGACTGTA 2942
 Db 3627 GAAGCTGCCAGCGGCTGACTGTA 3652

RESULT 5
 AAA59130
 ID AAA59130 standard; DNA; 5486 BP.
 XX
 AC AAA59130;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS2.
 XX
 KW Human; growth hormone inhibited tumour suppressor protein; hGHITS;

antineoplastic; dwarfism; gigantism; acromegaly; angiopathy; diabetic nephropathy; cardiopathy; tumour; breast cancer; renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 CDS 387..3653
 FT /*tag= a
 FT /*product= "tumour suppressor protein"
 XX
 PN EP1022333-A1.
 XX
 PD 26-JUL-2000.
 XX
 PF 07-OCT-1999; 99EP-0119199.
 XX
 PR 25-JAN-1999; 99JP-0016223.
 XX
 PA (JCRP-) JCR PHARM CO LTD.
 XX
 PI Koga J, Kono K, Zolotaryov FN;
 XX
 DR WPI; 2000-516013/47.
 XX
 PS P-PSDB; AAB07664.
 XX
 CC New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 CC diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 CC nephropathy or cardiopathy
 CC
 CC Claim 1; Page 28-39; 59pp; English.
 CC
 CC The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnosis of dwarfism, gigantism, acromegaly, angiopathy, or in malignant
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.
 CC
 SQ Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;

Query Match 55.4%; Score 1748; DB 21; Length 5486;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 2388; Conservative 0; Mismatches 545; Indels 153; Gaps 16;
 QY 1 ATGAGAGCACCCCAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCGATATTC 60
 Db 576 ATGAGAGCACCCCAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCGATATTC 635
 QY 61 CTCCTGCTTTTGGCAACAGTACAGGCTTTCGGAGCTGACAGGTGAACCGGAGATG 120
 Db 636 TTGTTGCTTTTGGCAACAGTACAGGCTTTCGGAGCTGACAGGTGAACCGGAGATG 695
 QY 121 CTTAGGAGTGGTGAATGCGGCATGTACACAGGAGATGGTGGCAGAGCGCTACGCAG 180
 Db 696 CTGACGAGTGGTGAATGCGGCATGTACACAGGAGATGGTGGCAGAGCGCTACGCAG 755
 QY 181 ACGGCGAGTGGATGATCAAGCTGCTTGGAGTACATCAGTAAGATGGCTACCTGGAC 240
 Db 756 ACTGGCAGCAGGATGATGAGGCGCTTGGAGTACATCAGTAAGATGGCTACCTGGAC 815
 QY 241 CCCAGGAATGAGCAGATTTGCGAGTCTATCAAGCAGAGCTCCCGAGGAAGGCGCTGCG 300
 Db 816 CCGAGGAATGAGCAGATTTGCGAGTCTATCAAGCAGAGCTCCCGAGGAAGGCGCTCATG 875
 QY 301 TCCACCGCGGTGACTGCGGCGGCGGCTTTCGAGGAGCAGGAGGAGCACTCCATCCAC 360
 Db 876 CCAACCGCGGTGACTGCGGCGGCGGCTTTCGAGGAGCAGGAGGAGCACTCCATCCAC 935

QY	361	CACCA	AGTGGTGGTGTG	GAAACTACAG-	-----GCCCCCGCGCA	399				
Db	936	CACCA	AGCTAGCGGT	TACCCCTACGAGGGCC	CAAGCTTCGCGCGT	GACGGCCCGACGGCG	995			
QY	400	CTG	AGGAGATCC	CGCGCAATATTTAG	ACTTCTCTTCCCT	GGAGCCGAGCGCGCACC	459			
Db	996	CTGG	AGGAGATGCC	CGGCGCGT	AGCTGGACTTACCTTTTCCCGGAGTC	GGCCCCCACGGG	1055			
QY	460	CAC	GTGCCAGGGT	TCA	CCAGCATCTCC	CAAGGGTACAGCAGCAG	TAGAGCCAA	519		
Db	1056	CCG	GCCACCAG-	---CACCAG	CCCCAACAGGCTAC	GGTCCAGCGT	TAGAGCAGCA	1112		
QY	520	GCG	-----	CAC	TTCCGGGCACAC	ATATGGTGGTGGT	CACTACT	564		
Db	1113	GGG	GCACACTT	CCCGCTCG	AGGGGCGCACTTAC	CGGGCGCGCACTGCT	TGGTGCCTGGG	1172		
QY	565	GAG	CAGTCTGG	TATGGGTG	CAGCGAGTTTC	CTTCCTCAGAAC	AGAGCGCCAC	624		
Db	1173	GAA	CCCTGGGCT	ACG	GAGTGCAGCGCCCTCTCTTCC	CAGAGCAGCGCCGCGGAG	1232			
QY	625	GCC	-----	TAT	TCAGCATGGCC	---AAGSCC	AGGGTGGCCCTCC	672		
Db	1233	ACCG	GGGTTAC	CGCCAGCCTGCC	CACGAAGGCGC	AGGAGACCGCAG	CGCGCCCTC	1292		
QY	673	AC	TTTCTG	CCCCATGCTG	GGCTGTAC	ACTGCTCGCACC	ACAAAGCGGGCGGTAC	732		
Db	1293	GCT	TTTCCACCCCT	TGCGCGGGCTCT	AGTGC	CGCACCCAC	CACCAAGCAGGCGCGT	1352		
QY	733	CTG	GGGCCCCAC	CAATAT	ATGTGTGGGAC	ACCGGGGTCCC	CAGTTACTGCGCAAA	792		
Db	1353	CCG	TGGCCCCAC	CACTGCTG	GGCTCCG	CAGCAGCGTGT	TCCGCCAGCAGC	1412		
QY	793	TC	GCACAGG	TGTGCTGG	CACCGTGC	AGAACCGCTCA	ATGCTGACTGTGT	TACGAGCTG	852	
Db	1413	CCCC	CGCAGAGC	CTCTCA	CTCCCTCG	CGAAACAGCCTCA	ACGTGGACCTGTAT	GAATTG	1472	
QY	853	GG	-----	CTC	ACGGTGCCTGGTCT	TCAGCTCC	ACTGGCAGCGCG	GNACTCGCTGCAG	906	
Db	1473	AGC	AGACCTCC	GTCCAG	ATGGCGCGGTG	CCACCTG	CGCGCGCGGAGCTCC	CTGTCAG	1532	
QY	907	AAG	CAGGCT	TAGAAG	CTCGCGC	---CG	CATGTGCTTTT	CGGGTGGC	954	
Db	1533	AAG	CGCGCCT	TGGAGG	CGCGCGCGC	AGTGGCTTCC	GGCCTGACTGCC	CAGTG	1592	
QY	955	CCC	ACAGAC	CACTCTT	CAACAAC	CCACAAC	-----CT	990		
Db	1593	CCC	ACAGAGCA	CACTCTT	CAACAGGCA	CCAGCGCGCGCGT	CGCGTGGCAAGGC	1652		
QY	991	GAG	CTC	ACTGCCCG	CCCCCAAC	ACAGGTAC	CGCGT	AGCGCGCACACATCTT	CA	1050
Db	1653	GAG	CTCCCTCG	CGCGCCCC	CAACACCGT	GACGCTGT	CACGGCGCGCACATCTT	GGCAC	1712	
QY	1051	CT	GTGAAG	AGCGTGC	GTGCTCGCGCGG	CCCCAGAC	AGCCGTGG	GGCGCCTCGCAC	1110	
Db	1713	CGG	TGAAG	AGCGTGC	TGAGG	CGGAGCGC	AGAGCGTGTGG	GGCCCTCGCAC	1772	
QY	1111	CCC	GCTGGTGG	CTGGC	-----	CC	ACAGCACTGCC	ACTGAG	1149	
Db	1773	CCG	CGCTGGG	TGCGCGCGCT	TGCCCCCG	CCCCCGCGCGCGCGCT	CGCGCGCGCTGGGAG	1832		
QY	1150	AGC	CTGGAG	CAAGGA	-----	GGG	CAGCGCGCCAC	ACCGCTG	1191	
Db	1833	GCT	TGGACG	CAAGGAGG	ACATGCC	CTGGCGTGG	GGCGCGAGGGCGT	TCCGCTG	1892	
QY	1192	GAT	GTG	GA	CTATGG	CGGCTCCG	AGCGAGGTGCC	ACCGCTCCG	TATCCAA	1251
Db	1893	GAC	GTGG	ATACG	GAGGCC	ACAGCCG	GAGTGGCG	CTCCGCC	TACCG	1952
QY	1252	CTG	CTGCC	CACTA	GTGAC	AGTAC	AGCGTGG	ACCTGG	ACAGCTGTG	1311
Db	1953	CTG	CTGCC	CAAGTGG	AGCAGTAC	-----	GACT	TGGACGCTGTG	CGCAGCATG	2006

1312	QY	CAGCAGAGTCTCCGAGGGG---CACTCATCTACACGGGAGTGCACAGACGCCACAAGGT	1366
2007	Db	GAGCAGAGGCTCCGTGCGGGCCCCAACGAGCCGAGGGCGCGCAAGAGCCCAAGC	2066
1369	QY	GCGAAGGAGACAAAGCTGGCAGAGACAAAAGCAGATTACAGACCTCCCGGTGCCTGTC	1428
2067	Db	GCCAAAGGGGACAAAGCGGAAAGGATAAAAGCAGATTACAGACCTCCCGTTCGCGTC	2126
1429	QY	CGCAAGATATGACAGAGATGAAGAGAAGAGAGTCTCGCATCAAGAGTTACTCCCTTAT	1488
2127	Db	CGCAAAACAGCAGAGACGAGAGAAGAGAGTCAACGCATCAAGAGCTACTCGCCATAC	2186
1489	QY	GCGTTCAAATCTTCATGGAGACAACGCTGGAGAATGTCAATAAACTTACCAGCAGAAG	1548
2187	Db	GCGTTTAACTTCTTCATGGAGCAGCAGCTGGAGAAATGTCAATAAACTTACCAGCAGAAG	2246
1549	QY	GTACCCGAGGCTACAGCTGGAGCAGGAATGCCAAAGCTGGGCTCTGTGAGGCCGAG	1608
2247	Db	GTTAACGGAGGCTCGAGCTGGAGCAAGAAATGCCAAAGCTGGAGCTCTGTGAAGCTGAG	2306
1609	QY	CAGGAGCAGATCAGGAAGATCCTCTTACCAGAAGAGGTCTAACTACAACCGGCTGAAGAGG	1668
2307	Db	CAGGAGCAGATCGGAAGATCCTCTTACCAGAAGAGGTCTAACTACAACAGGTTAAGAGG	2366
1669	QY	GCCAGATGGACAAGTCCATGTTGTGAAAATCAAGACTCTAGGCAATCGGCTTTGGG	1728
2367	Db	GCCAAATGGACAAGTCTATGTTGTCAAGATCAAAACCCCTGGGGATCGGCTTTGGA	2426
1729	QY	GAAGTGTGCTCGCTGTGAAGCTGGACACTCAAGCTCTGTACGCCATGAAGACTCTCAGG	1788
2427	Db	GAAGTGTGCTTGTGTGAAGTGGACACTCACGCCCTGTACGCCATGAAGACCCTAAGG	2486
1789	QY	AAGAAGATGTCTTCAACCGGAATCAAGCTGGCCCATGTCAAGGCTGAGAGGGACATCCTTG	1848
2487	Db	AAAAAGGATGTCTTCAACCGGAATCAAGCTGGCCCATGTACGCCATGAAGACCCTAAGG	2546
1849	QY	GCTGAAGCAGACAATGAGTGGGTGGTCAAACCTCTACTCTCTTCCAGGACAAAGACAGC	1908
2547	Db	GCCGAGGCAGACAATGAGTGGGTGGTCAAACCTCTACTCTCTTCCAAAGACAAAGACAGC	2606
1909	QY	CTGTACTTTGTATGGACTACATACCAGGCGGGATATGATGAGCCTGTGTATCAGGATG	1968
2607	Db	CTGTACTTTGTATGGACTACATCCCTGGTGGGACATGATGAGCCTGTGTATCAGGATG	2666
1969	QY	GAGGTCTTCCCTGAGCACTGGCCGGCTTCTACATTCGAGAGTTGACCCCTGGCCATTGAA	2028
2667	Db	GAGGTCTTCCCTGAGCACTGGCCGGCTTCTACATTCGAGAGTGTACTTTGGCCATTGAG	2726
2029	QY	AGTGTCCAAGATGGGCTTTTATCCACGGGACATCAAGCCTGCACAACTACTATCTCAG	2088
2727	Db	AGTGTCCAAGATGGGCTTATCCACGGGACATCAAGCCTGCATACATTTTGTATAGAT	2786
2089	QY	CTGGATGGTCATATTAAAGCTGACAGATTTTGGCCTCTGCACTGGATTCAGTGGGACTCAC	2148
2787	Db	CTGGATGGTCATATTAAACTCACAGATTTTGGCCTCTGCACTGGGTTCAAGTGGGACTCAC	2846
2149	QY	AATTCCAGTACTACCAGAAAGGAACACATGACAGGACAGCATGAGAGCCCGGTGAC	2208
2847	Db	AATTCCAAATATTACCAGAAAGGAGCCATGTCAGACGGACAGCATGAGAGCCCGCAGC	2906
2209	QY	CTCTGGGCGATGTTTCCAACCTGTGCTGTGGAGCAGAGTTAAAGACCCCTGGAGCAGAGG	2268
2907	Db	CTCTGGGATGATGCTCTTAAGTGTGGGTGGGACAGGCTGAAGACCCCTAGAGCAGAGG	2966
2269	QY	GGCAGAAAGCAGCACAGAGTGGCTGGCACAATTCTCTGTGGGACACCAAAATTACATC	2328
2967	Db	GGCGGGAAGCAGCACAGAGTGGCTGGCACAATTCTCTGTGGGAGGAGTCCCAACTACATC	3026
2329	QY	GCTCCGGAGGTGCTTCTCCGCAAAAGGTTACAGCAGCTCTGTGACTGGTGGAGCGCTCGGT	2388
3027	Db	GCACCCGAGGTTGCTCCTCCGCAAAAGGTTACACTCAACTCTGTGACTGGTGGAGTGTGGA	3086
2389	QY	GTGATTCTCTTTGAGATGCTGTTGGGCAGCGCGCTTTCTTTGGCCCCCACCACCCACAGAG	2448

CC gene Lats2. Lats2 is a kinase and can be used for gene diagnosis of
CC cancers.
XX
SQ Sequence 3533 BP; 831 A; 1096 C; 1011 G; 595 T; 0 other;
Query Match .54.4%; Score 1717.2; DB 22; Length 3533;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 2382; Conservative 0; Mismatches 548; Indels 156; Gaps

[illegible]

RESULT 7

AAZ61160
ID AAZ61160 standard; cDNA; 1961 BP.
AC AAZ61160;
XX 30-MAY-2000 (first entry)
DT cDNA S81771A encoding a partial protein kinase.
DE Kinase activity; molecular weight marker; isoelectric focusing marker;
KW peptide fragmentation control; cellular signal transduction; ss.
XX Homo sapiens;
FH Key
FT Location/Qualifiers
FT 3.1769
FT /*tag= a
FT /note= "partial sequence"
XX WO200008180-A2.
XX 17-FEB-2000.
XX 03-AUG-1999; 99WO-US17630.
XX 04-AUG-1998; 98US-0095270.
XX 11-SEP-1998; 98US-0099972.
XX (IMV) IMMUNEX CORP.
XX Virca GD, Bird TA, Anderson DM, Marken JS;
XX WPI; 2000-195584/17.
XX P-PSDB; AAY69162.
XX New human kinase polypeptides and polynucleotides used as molecular
XX weight markers and as controls for peptide fragmentation
XX Claim 1; Page 9; 60pp; English.
XX The present sequence encodes a partial polypeptide which has kinase
XX activity. The kinase polynucleotides can be used to express the
XX polypeptides, and as probes to identify nucleic acids encoding
XX proteins having kinase activity. The kinase polypeptides and
XX fragmented polypeptides are used as molecular weight and isoelectric
XX focusing markers, and as controls for peptide fragmentation. They also
XX have a number of therapeutic uses as kinases play a central role in
XX cellular signal transduction. The polypeptides could also be used to
XX identify binding partner proteins. The polypeptides can also be used as
XX a reagent to identify any proteins that the polypeptide regulates, and
XX proteins with which it might interact. The polypeptides may also be
XX used for preparation of antibodies. The antibodies can be used in
XX assays to detect the presence of the protein, and to purify the protein
XX by immunoaffinity chromatography.
XX Sequence 1961 BP; 504 A; 550 C; 539 G; 368 T; 0 other;
XX Query Match 41.0%; Score 1293; DB 21; Length 1961;
XX Best Local Similarity 84.4%; Pred. No. 0;
XX Matches 1497; Conservative 0; Mismatches 255; Indels 21; Gaps 3;
QY 1185 CCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGCCACCGCCTCCCGTATCCCAA 1244
DB 2 CCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGCCACCGCCTCCCGTATCCCAA 61
QY 1245 GCATCTGCTGCTGCCAGTAAGTCTGAGAGTACAGCGTGGACCTGGAGAGCTGTGCAC 1304
DB 62 GCACCTGCTGCTGCCAGTAAGTCTGAGAGTACAGCGTGGACCTGGAGAGCTGTGCAC 115
QY 1305 CAGTGTGACAGAGTCTCCGAGGGG---CACTGATCTAGACGGGAGTGAACAGGCA 1361
DB 116 AGGCATGGAGAGAGCTCCCTGGTGGGGGCCCCAACGAGCCCGGCGGCAAGAGCGG 175

QY 1362 CAAAGGTGCGAAGGGAGACAAAGCTGGCAGAGACAAAAGAGATTCAGACCTCCCGGT 1421
DB 176 CAAAGGTGCGAAGGGAGACAAAGCTGGCAGAGACAAAAGAGATTCAGACCTCCCGGT 235
QY 1422 GCCTGTCCGCAAGATAGCAGAGATGAGAGAGAGAGAGTCTCGATCAAGAGTACTC 1481
DB 236 TCCCGTCCGCAAGATAGCAGAGATGAGAGAGAGAGAGTCTCGATCAAGAGTACTC 295
QY 1482 CCCTTATGCTTCAATCTTCTCATGGAGCAACACAGTGGAGAGTGTATCAAAACCTACCA 1541
DB 296 GCATACGCGCTTAACTTCTCATGGAGCAGCAGTGGAGAGTGTATCAAAACCTACCA 355
QY 1542 GCAGAAGGTACGCGGAGGTACAGCTGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGA 1601
DB 356 GCAGAAGGTAAACCGGAGGTACAGCTGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGA 415
QY 1602 GCCGAGCAGGAGAGATGAGGAGATCTCTTACGAGAGAGAGTCTAAGTCAACCGGT 1661
DB 416 AGCTGAGCAGGAGAGATGCGGAGATCTCTTACGAGAGAGTCTAATTAACAGAGTT 475
QY 1662 GAAGAGGCCAAGATGGACAGTCCATCTTCTGAGAAATCAAGACTCTAGGCTCGGTGC 1721
DB 476 AAAGAGGCCAAGATGGACAGTCTATGTTTCAAGATCAAAACCTGGGATCGGTGC 535
QY 1722 CTTTGGGAAAGTGTGCTTGTAAAGCTGGACACTCAAGCTCTGTAGCCATGAAGAC 1781
DB 536 CTTTGGGAAAGTGTGCTTGTAAAGCTGGACACTCAAGCTCTGTAGCCATGAAGAC 595
QY 1782 TCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGA 1841
DB 596 CCTAAGGAAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGA 655
QY 1842 CATCTGCTGAGCAGACAAATGAGTGGTGTGCTCAAACTCTACTACTCTTCCAGAGCAA 1901
DB 656 CATCTGCTGAGCAGACAAATGAGTGGTGTGCTCAAACTCTACTACTCTTCCAGAGCAA 715
QY 1902 GGACAGCTGTACTTTGTGATGGACTACATACCGCGGGGATATGATGAGCTCTGAT 1961
DB 716 AGACAGCTGTACTTTGTGATGGACTACATACCGCGGGGATATGATGAGCTCTGAT 775
QY 1962 CAGGATGAGGTCTTCCCTGAGCAGCTGGCCGCTTCTACATTCAGAGTTGACCTGGC 2021
DB 776 CCGGATGAGGTCTTCCCTGAGCAGCTGGCCGCTTCTACATTCAGAGTTGACCTGGC 835
QY 2022 CATTAAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCTCTGACCAATCT 2081
DB 836 CATTAAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCTCTGACCAATCT 895
QY 2082 CATCAGCTGGATGTCTATTAAGTGAAGTGAAGTGTGGCTCTGCTGATTCAGGTG 2141
DB 896 GATAGATCTGGATGTCTACATTAAGTGAAGTGAAGTGTGGCTCTGCTGATTCAGGTG 955
QY 2142 GACTCACAATCCAGTACTACCAAGAGGAAACACATGAGACAGGACAGCTGAGCC 2201
DB 956 GACTCACAATCCAGTACTACCAAGAGGAAACACATGAGACAGGACAGCTGAGCC 1015
QY 2202 CGGTGACCTCTGGGAGAGTGTTCACACTGTCTGAGAGAGTGAACACCTGGA 2261
DB 1016 CAGCCAGCTCTGGGATGATGTCTTAAGTGTCTGAGAGAGTGAACACCTGGA 1075
QY 2262 GCAGAGGCGCAGAGCAGCAGAGGTGCTTCCGAAAGGTTACAGGAGCTGTGACTGGAG 2321
DB 1076 GCAGAGGCGCAGAGCAGCAGAGGTGCTTCCGAAAGGTTACAGGAGCTGTGACTGGAG 1135
QY 2322 TTACATCGCTCGGAGGTGCTTCCGAAAGGTTACAGGAGCTGTGACTGGAG 2381
DB 1136 TTACATCGCTCGGAGGTGCTTCCGAAAGGTTACAGGAGCTGTGACTGGAG 1195
QY 2382 CGTCTGCTGTCTCTTTGAGATGCTGTTGGGAGCGCTTCTTGGCCCGCCACCC 2441
DB 1196 TGTGAGTGTCTCTTTCGAGATGCTGTTGGGAGCGCTTCTTGGCCCGCCACCC 1255
QY 2442 CACAGAGAGCGAGCTGAAGGTGATCAACTGGGAGAGAGCTGATATCTCTACGAGT 2501

Db 1256 CACAAACCCAGCTGAAGGTGATCAACTGGGAGAACACGCTCCACATTCAGGCCAGGT 1315
 Qy 2502 GAGGCTCAGCGCTGAGGCGGAGAGCTCATCAGCAAGCTGTGCTGGCGGCTGACTGCGG 2561
 Db 1316 GAAGCTGAGCCCTGAGGCGGAGAGCTCATCAGCAAGCTGTGCTGGCGGAGAGCCAG 1375
 Qy 2562 CTTGGGAGGATGGGCGGAGAGCTCATCAGCAAGCTGTGCTGGCGGAGAGCCAGT 2621
 Db 1376 CTTGGGAGGATGGGCGGAGAGCTCATCAGCAAGCTGTGCTGGCGGAGAGCCAGT 1435
 Qy 2622 TTCCCGTACATCCGAAGAGGCTGACCTAGTCCCGACCATCAGCCACCCATGGA 2681
 Db 1436 CTCAGTGACATCCGGAAGAGCTGACCTAGTCCCGACCATCAGCCACCCATGGA 1495
 Qy 2682 CACCTCCAAATTTGACCGGCTGGATGAAGAAAGCCCTGGCAGGCGGAGAGAG 2741
 Db 1496 CACCTCGAATTCGACCCCTGATGATGAAGAAAGCCCTGGCAGGCGGAGAGAG 1555
 Qy 2742 CGCAAGGCTGGGAGAGCTGGGCTCCCGACGAGAGAGAGAGAGAGAGAGAGAG 2801
 Db 1556 CACCAAGGCTGGGAGAGAGCTGGGCTCCCGACGAGAGAGAGAGAGAGAGAGAG 1615
 Qy 2802 TGAGTTACCTTCGCGAGGCTTCGATGACAGCGCTATCCCTTCGCGGCGGAGAG 2861
 Db 1616 CGAATTCACCTTCGCGAGGCTTCGATGACAGCGCTATCCCTTCGCGGCGGAGAG 1675
 Qy 2862 CTCAG-----AGCCCGCAGAGAGTGCAGAGCCAGGAGTCCGAGTTCGAGAG 2909
 Db 1676 TTCAGGAGCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTGTGATGTTGA 1735
 Qy 2910 TCGCGCGAGGCTGCGAGCGGCTGAGAGCTCAGATTTAGAAAGCTGTGATGTTGA 2942
 Db 1736 TCAGACTGAAGCTGCCAGCGCTGTGATGTTGA 1768

RESULT 8
 AA261158
 ID AA261158 standard; cDNA; 1498 BP.
 XX AC AA261158;
 XX 30-MAY-2000 (first entry)
 DE cDNA SS1771 encoding a human protein kinase.
 XX Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FT CDS 2..1498
 FT /tag= a
 FT /product= "kinase"
 XX
 PN W0200008180-A2.
 XX 17-FEB-2000.
 XX
 XX 03-AUG-1999; 99WO-US17630.
 XX
 XX 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0099972.
 XX
 XX (IMV) IMMUNEX CORP.
 XX
 XX Virca GD; Bird TA, Anderson DM, Marken JS;
 XX WPI; 2000-195584/17.
 DR P-PSDB; AAY69160.
 XX
 XX New human kinase polypeptides and polynucleotides used as molecular

PT weight markers and as controls for peptide fragmentation
 XX Claim 1; Page 8-9; 60pp; English.
 XX The present sequence encodes a polypeptide which has kinase
 CC activity. The kinase polynucleotides can be used to express the
 CC polypeptides, and as probes to identify nucleic acids encoding
 CC proteins having kinase activity. The kinase polypeptides and
 CC fragmented polypeptides are used as molecular weight and isoelectric
 CC focusing markers, and as controls for peptide fragmentation. They also
 CC have a number of therapeutic uses as kinases play a central role in
 CC cellular signal transduction. The polypeptides could also be used to
 CC identify binding partner proteins. The polypeptides can also be used as
 CC a reagent to identify any proteins that the polypeptides regulate, and
 CC proteins with which it might interact. The polypeptides may also be
 CC used for preparation of antibodies. The antibodies can be used in
 CC assays to detect the presence of the protein, and to purify the protein
 CC by immunofluorescence chromatography.
 XX
 SQ Sequence 1498 BP; 384 A; 417 C; 427 G; 270 T; 0 other;
 Query Match 36.4%; Score 1148.4; DB 21; Length 1498;
 Best Local Similarity 86.3%; Pred No. 3.5e-288;
 Matches 1295; Conservative 0; Mismatches 196; Indels 9; Gaps 2;
 QY 1185 CCGCTGGATGTGACTATGCGGCTCCGAGCGGAGTGCACCGCTCCGCTATCCAAA 1244
 Db 4 CCGCTGGATGTGACTATGCGGCTCCGAGCGGAGTGCACCGCTCCGCTATCCAAA 63
 QY 1245 GCATCTGCTGCTGCCAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1304
 Db 64 GCATCTGCTGCTGCCAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 117
 QY 1305 CAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1361
 Db 118 AGCATGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 177
 QY 1362 CAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1421
 Db 178 CAAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 237
 QY 1422 GCCTGTGCTGCTGCCAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1481
 Db 238 TCCGCTGCCAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 297
 QY 1482 CCCTATGCTTCAAAATTTCTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1541
 Db 298 GCATAGCTTTAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 357
 QY 1542 GCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1601
 Db 358 GCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 417
 QY 1602 GCGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1661
 Db 418 AGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 477
 QY 1662 GAAGAGGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1721
 Db 478 AAAGAGGCTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 537
 QY 1722 CTTTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1781
 Db 538 CTTTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 597
 QY 1782 TCATCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1841
 Db 598 CTTAAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 657
 QY 1842 CATCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1901
 Db 658 CATCTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 717

Qy 1902 GGACAGCTGTACTTTGTGATGACTACATACAGCGGGGATATGATGAGCTGCTGAT 1961
 Db 718 AGACAGCTGTACTTTGTGATGAGCTACATCCTGGTGGGACATGATGAGCTGCTGAT 777
 Qy 1962 CAGGATGAGGTCCTCCCTGAGCAGCTGGCCGGTCTTACATTTGAGAGTTGACCTGGC 2021
 Db 778 CCGGATGAGGTCCTCCCTGAGCAGCTGGCCGGTCTTACATTCGAGAGCTGACTTTGGC 837
 Qy 2022 CATTGAAGAGTCCACAGATGGGCTTTATCCACGGGACATCAAGCTGACACATACT 2081
 Db 838 CATTGAGAGTCCACAGATGGGCTTTATCCACGGGACATCAAGCTGATTAACATTTT 897
 Qy 2082 CATCCAGCTGGATGCTATATTAAGCTGACAGATTTTGGCCCTCTGCACTGGATTGAGTG 2141
 Db 898 GATAGATCTGGATGCTACATTAAGCTGACAGATTTTGGCCCTCTGCACTGGATTGAGTG 957
 Qy 2142 GACTCACATTTCCAAAGTACTACAGAAAGGAAACATGAGACAGGACATGGAGCC 2201
 Db 958 GACTCACATTTCCAAATATTAACAGAAAGGAGCCATGTCAGACAGGACATGGAGCC 1017
 Qy 2202 CGGTGACCTCTGGGACGATTTTCAACTGTGCTGTGGACAGGTTAAAGACCTTGA 2261
 Db 1018 CAGCGACCTCTGGGATGATGTCTTAAGTGTGCTGTGGGACAGGCTGAAGACCTTGA 1077
 Qy 2262 GCAGAGGGCGGAGAGCAGCAGCAGGTCCTGGCAGATTTCTGTGCGGACACCAA 2321
 Db 1078 GCAGAGGGCGGAGAGCAGCAGGTCCTGGCAGATTTCTGTGCGGACACCAA 1137
 Qy 2322 TTACATGCTCGGAGGTCCTTCCGCAAGGGTACAGCAGCTCTGTGACTGTGGAG 2381
 Db 1138 CTACATGCTCGGAGGTCCTTCCGCAAGGGTACAGCAGCTCTGTGACTGTGGAG 1197
 Qy 2382 CGTGGTGTGATCTCTTTGAGATGCTGTGGGACGCGCTTTTGGCCCGCCACCC 2441
 Db 1198 TGTGGTGTGATCTCTTTGAGATGCTGTGGGACGCGCTTTTGGCCCGCCACCC 1257
 Qy 2442 CACAGAGCGGAGTGAAGGTGATCAACTGGGAGAGCAGCTGTCATATCCCTACGAGT 2501
 Db 1258 CACAGAGCGGAGTGAAGGTGATCAACTGGGAGAGCAGCTGTCATATCCAGCCAGT 1317
 Qy 2502 GAGGTCAGGCTGAGGCGGAGCCTCATCAGAGTGTGCTGCGCGGCTGAGTGGCG 2561
 Db 1318 GAACTGAGGCTGAGGCGGAGCCTCATCAGAGTGTGCTGCTGCGGAGAGCAGCG 1377
 Qy 2562 CCTGGGAGGATGGGAGGATGACCTCAAGCAGCAGCTTCTTCAACACATGACTT 2621
 Db 1378 CTTGGGCGGATGGGCGGATGACCTGAAGCCCGCCCTTTCTTTCAGGCGCATTTGACTT 1437
 Qy 2622 TTCCGTCGATCCGAAAGCAGGCTGACCTTACCTCCCGCCAGCCAGCCCGCATGGA 2681
 Db 1438 CTCCAGTCATCCGAAAGCAGCAGCCCTTACCTTCCAGCCAGCCCGCATGGA 1497

RESULT 9

ABK43457
 ID ABK43457 standard; cDNA: 1912 BP.

XX AC ABK43457;

XX DT 05-JUN-2002 (first entry)

XX DE DNA encoding novel central nervous system protein #37.

XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angiodysplasia;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy; gene; ss.

XX OS Homo sapiens.
 XX PN WO200155318-A2.
 XX PD 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01332.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.

PA (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-581633/65.
XX P-PSDB; AAU87127.
DR New isolated nucleic acid encoding a protein for diagnosing,
XX preventing, treating or ameliorating medical conditions and used as
PT food additives or preservatives
PT Claim 1; SEQ ID No 47; 837pp; English.
XX The invention describes an isolated nucleic acid molecule (I) encoding a
XX novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical condition and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC adenocarcinomas and irritable bowel syndrome, endocrine disorders e.g. diabetes
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
CC Query Match 35.5%; Score 1118.6; DB 23; Length 1912;
Best Local Similarity 85.9%; Pred. No. 2.2e-280;
Matches 1257; Conservative
Yy 1492 TTCAAATTTTCATGGAGCAACACGCTGGAGATGTCATCAAAACCTACAGCAGAGGTC 1551
Db 1 TTTAAGTCTTTCATGGAGCAGCAGCTGGAGATGTCATCAAAACCTACAGCAGAGGTT 60
Yy 1552 AGCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTAGCCGAGCAG 1611
Db 61 AACCGAGGCTGCAGCTGGAGCAAGAAATGGCCAAAGCTGGCTCTGTAGCAGCAG 120
Yy 1612 GAGCAGATGAGGAGAGATCCTCTACCAAGAGGAGTCTAACTACAAACCGGTGAAGAGGCC 1671
Db 121 GAGCAGATGCGAAGATCCTCTACCAAGAGAGTCTAACTACAAACCGGTGAAGAGGCC 180
Yy 1672 AAGATGGACAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGTGGCTTTGGGAA 1731
Db 181 AAGATGGACAGTCTATGTTGTCAAGATCAAAACCGGTGGATCGGCTTTGGAGAA 240
Yy 1732 GTGTGCTTGTGTAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAG 1791
Db 241 GTGTGCTTGTGTAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAG 300
Yy 1792 AAGATGTCCTGAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGAGCATCTGGCT 1851
Db 301 AAGATGTCCTGAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGAGCATCTGGCT 360
Yy 1852 GAAGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGAGCAGAGCAGCTG 1911
Db 361 GAGGAGACAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGAGCAGAGCAGCTG 420
Yy 1912 TACTTTGTGATGACTACATACAGCGGGGATATGATGAGCCTGCTGATCAGATGAG 1971
Db 421 TACTTTGTGATGACTACATACAGCGGGGATATGATGAGCCTGCTGATCAGATGAG 480

QY 1972 GTCTCCCTGAGCACCTGCGCCCTTCTACATTGTCAGAGTGTACCTGCGCCATTGAAGT 2031
 Db 481 GTCTCCCTGAGCACCTGCGCCCTTCTACATTGTCAGAGTGTACCTGCGCCATTGAAGT 540
 QY 2032 GTCCACAAGTGGCTTTTTCACCCGGGACATCAAGCCTGACAACATCATCTGACCGT 2091
 Db 541 GTCCACAAGTGGCTTTTTCACCCGGGACATCAAGCCTGACAACATTTGATAGATCG 600
 QY 2092 GATGGTCATTAAGTCAGACATTTTGGCTCTGCACTGGATTAGTGGACTCAAT 2151
 Db 601 GATGGTCATTAAGTCAGACATTTTGGCTCTGCACTGGATTAGTGGACTCAAT 660
 QY 2152 TCCAGTACTACAGAAAGGACACATGAGACAGACAGATGAGAGCCCGGTGACCTC 2211
 Db 661 TCCAGTACTACAGAAAGGACACATGAGACAGACAGATGAGAGCCCGGTGACCTC 720
 QY 2212 TGGGAGATGTTTCCAACTGTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCG 2271
 Db 721 TGGGAGATGTTTCCAACTGTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCG 780
 QY 2272 CAGAGCAGCAGAGAGTGCCTGGCACATTTCTTGTGGGACACCAATTTACATCGT 2331
 Db 781 CAGAGCAGCAGAGAGTGCCTGGCACATTTCTTGTGGGACACCAATTTACATCGT 840
 QY 2332 CCGGAGGTGCTTCTCCGCAAGGTTACACGAGCTGTGACTGTGGAGCGTGGTGTG 2391
 Db 841 CCGGAGGTGCTTCTCCGCAAGGTTACACGAGCTGTGACTGTGGAGCGTGGTGTG 900
 QY 2392 ATTCTCTTTGAGATGCTGTGGAGCCGCTTTCTTGGCCCGCCACCCACACAGAGCG 2451
 Db 901 ATTCTCTTTGAGATGCTGTGGAGCCGCTTTCTTGGCCCGCCACCCACACAGAGCG 960
 QY 2452 CAGCTGAAGTGATCAACTGAGGAGCAGCTGTCATATCCCTACGAGTGGGCTCAGC 2511
 Db 961 CAGCTGAAGTGATCAACTGAGGAGCAGCTGTCATATCCCTACGAGTGGGCTCAGC 1020
 QY 2512 GCTGAGGCCGAGACCTCATCAAGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGCGG 2571
 Db 1021 GCTGAGGCCGAGACCTCATCAAGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGCGG 1080
 QY 2572 GATGGGCGAGATGACCTCAAGGACACACCGTCTTCTCAACACATCGACTTTTCCCGTGC 2631
 Db 1081 GATGGGCGAGATGACCTCAAGGACACACCGTCTTCTCAACACATCGACTTTTCCCGTGC 1140
 QY 2632 ATCCGAAGCAGCTGACCTTACCTCCACCATCCAGCAGCCACCCATGACACCTCCAT 2691
 Db 1141 ATCCGAAGCAGCTGACCTTACCTCCACCATCCAGCAGCCACCCATGACACCTCCAT 1200
 QY 2692 TTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGCCGAGGAGAGCGCCAGGCG 2751
 Db 1201 TTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGCCGAGGAGAGCGCCAGGCG 1260
 QY 2752 TGGGACACCTGCTTCCCGGAGCAGACAGATCCAGACAGCGCTTCTATGAGTTACG 2811
 Db 1261 TGGGACACCTGCTTCCCGGAGCAGACAGATCCAGACAGCGCTTCTATGAGTTACG 1320
 QY 2812 TTTCCGAGGTTCTGATCACAAGGCTATCCCTTCCCGTGGCCGAGCCCTCAG - - - - 2866
 Db 1321 TTTCCGAGGTTCTGATCACAAGGCTATCCCTTCCCGTGGCCGAGCCCTCAG - - - - 1380
 QY 2867 - - - - - AGCCCGCAGAGAGTGCACACCGAGGGGATGCGGACTTGAAGGTCGCGCCGAG 2919
 Db 1381 GAAGCTTACAGGCTGAGAGTGCATTTTGAAGGCTGATCTGATCTGATGATCAGACTGAA 1440
 QY 2920 GCTGCGCAGCGGTGACCTGTA 2942
 Db 1441 GCTGCGCAGCGGTGACCTGTA 1463

RESULT 10
 AAD03989
 ID AAD03989 standard; cDNA; 2043 BP.
 XX

AC AAD03989;
 XX 02-JUL-2001 (first entry)
 DT Human protein tyrosine kinase receptor (PTK) cDNA from clone HMUBT71.
 DE
 XX
 KW Human: protein tyrosine kinase receptor; PTK; antibacterial; antiviral;
 KW antifungal; antiparasitic; sarcoidosis; inflammation; asthma; arthritis;
 KW cardiovascular disorder; severe combined immunodeficiency; SCID; vaccine;
 KW hyperproliferative disorder; atherosclerosis; brain disorder; leukaemia;
 KW multiple sclerosis; Alzheimer's disease; vascular dementia; thrombosis;
 KW neurodegenerative disorder; gastrointestinal disorder; cancer; lymphoma;
 KW blood coagulation disorder; trauma; cerebrovascular disorder; tendinitis;
 KW gene therapy; nephrotic syndrome; glomerulonephritis; allergy; neoplasm;
 KW musculo-skeletal disorder; Parkinson's disease; autoimmune disorder;
 KW behavioural disorder; renal disorder; ss.
 XX Homo sapiens.
 OS
 XX
 FH Location/Qualifiers
 FT CDS
 FT 248..1585
 FT /*tag= a
 FT /product= "Human PTK protein from clone HMUBT71"
 XX
 XX WO200129564-A1.
 XX 26-APR-2001.
 XX 12-OCT-2000; 2000WO-US28066.
 PF 15-OCT-1999; 99US-0159542.
 PR 17-NOV-1999; 99US-0165914.
 PR 14-MAR-2000; 2000US-0189027.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Ruben SM, Shi Y, Young PE, Ni J;
 XX WPI: 2001-290976/30.
 DR P-PSDB; AAE00663.
 PT New isolated nucleic acid molecule for producing human protein tyrosine
 PT kinase receptor
 XX Claim 1; Page 260-261; 288pp; English.
 PS The present cDNA sequence encodes human protein tyrosine kinase receptor
 CC (PTK) from clone HMUBT71.
 CC The invention relates to human protein tyrosine kinase (PTK) receptors
 CC and their corresponding cDNA molecules. PTK receptors are useful in
 CC providing immunological probes for differential identification of the
 CC tissues or cell types present in a biological sample. PTK is used in
 CC methods for the diagnosis, prevention and treatment of various disorders
 CC related to PTK such as immune system disorders (severe combined
 CC immunodeficiency (SCID), inflammation); hyperproliferative disorders
 CC (neoplasm, sarcoidosis); cardiovascular disorders (arrhythmia,
 CC atherosclerosis); central nervous system disorders (multiple sclerosis),
 CC neurodegenerative disorders (Alzheimer's disease, Parkinson's disease);
 CC blood coagulation disorders (thrombocytopenia); autoimmune disorders
 CC (biliary cirrhosis, Crohn's disease); respiratory disorders (asthma,
 CC allergy); gastrointestinal disorders (inflammatory bowel disease);
 CC cerebrovascular disorders (thrombosis, vascular dementia); brain
 CC disorders (phenylketonuria); cancers (such as ovarian, lung, bladder,
 CC liver, breast and lymphomas); behavioural disorders (Pourette's
 CC syndrome); musculo-skeletal disorders (arthritis, trauma, tendinitis);
 CC renal disorders (nephrotic syndrome, glomerulonephritis); metastases of
 CC malignancies and related disorders (leukaemia, multiple myeloma); and
 CC infections caused by bacteria, viruses, fungi and parasites. PTK is also
 CC useful for screening therapeutic compounds. PTK is used as an antigen in
 CC a vaccine to raise an immune response against infectious disease. PTK
 CC nucleic acids are useful in gene therapy.
 XX Sequence 2043 BP; 551 A; 528 C; 495 G; 469 T; 0 other;

[illegible]

QY	2605	TTCAACACCATCGACTTTTCCCGTCGACATCCGAAAGCAGGCTGCACCCCTACGTCCTCCAC	2604		
Db	1235	TTTTCAGCGCCATGACTTCCAGTCGACATCCGGAAGCAGCAGCCCTACGTTTCCAC	1294		
QY	2665	ATCAGCGACCCCATGGACACCTTCCAATTTGACCCGGTGGATGAAGAAAGCCCTGGCAC	2724		
Db	1295	ATCAGCGCCATCGACATCGACACCTCGAATTTCCAGCCCGCTAGATGAGAAAGCCCTTGGAA	1354		
QY	2725	GAGGCCAGCGGAGAGCGCCCAAGCCCTGGGACACGCTGGGACACGCTGGCCCTCCCCAGCAGCAAGCAT	2784		
Db	1355	GATGCCAGCGAAGGTAGCACCAGCGCTGGGACACACTCACTCGCCCAATAACAAGCAT	1414		
QY	2795	CCAGACGACGCTTCTATGAGTTTCACTTCCGCGAGGTTCTTCCGATGACACGCGTATCCC	2844		
Db	1415	CCTGAGCAGCATTTTACGAATTCACCTTCGGAAGGTTCTTTGATGACAATGCTATCCCC	1474		
QY	2845	TTCCGGTCCCGAAGCCCTCAG-----AGCCCGCAGAGAGTGCGACACCCAGGG	2892		
Db	1475	TTTCGATGCCCAAGAGCTTCAGGACGAGAAGTTCAAGGCTCAGAGCTCAGATTTAGAA	1534		
QY	2893	GATCGGACTTGGGAAGGTGCGCGGAGGCTGCCAGCCGGTGACGTGTA	2942		
Db	1535	AGCTCTGATCTGGTGATCAGATGAAGGCTGCCAGCTGTGTACGTGTA	1584		
RESULT 11					
AAZ51506					
ID	AAZ51506 standard; DNA; 3213 BP.				
XX					
AC	AAZ51506;				
XX					
DT	21-JUN-2000 (first entry)				
XX					
DE	Mouse Lats (large tumour suppressor) DNA.				
XX					
KW	Mouse; Lats; large tumour suppressor; cytostatic; vulnery;				
KW	cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;				
KW	treatment; prevention; screening; cancer; skin; ovarian tumour;				
KW	soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;				
KW	LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;				
KW	dysplasia; degenerative disorder; growth deficiency; physical trauma;				
KW	hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.				
XX					
OS	Mus musculus.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	1..2889			
FT		/**tag= a			
FT		/product= "Lats protein"			
FT		/partial			
XX					
PN	WO200010602-A1.				
XX					
PD	02-MAR-2000.				
XX					
PF	18-AUG-1999; 99WO-US19068.				
XX					
PR	18-AUG-1998; 98US-0096996.				
PR	18-AUG-1998; 98US-0096997.				
XX					
XX	(UYUA) UNIV YALE.				
XX					
PI	Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;				
PI	Turenchalk GS, Stewart RA;				
XX					
XX	WPI: 2000-246496/21.				
DR	P-PSDB; AA70391.				
XX					
PT	Use of lats proteins, complexes of lats and cdc2 for treating cancer				
PT	that is refractory to treatment by standard chemotherapy and radiation				
PT	therapy, and disorders associated with aberrant levels of cdc2 activity				
XX					

Claim 44; Fig 13; 134pp: English.

PS The present sequence is a DNA encoding mouse Lats (large tumour
 XX suppressor) protein which is a cell overproliferation inhibitor and a
 CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 CC The present sequence is useful for treating cancer that is refractory
 CC to standard chemotherapy or radiation therapy such as hyperplasia,
 CC metaplasia, or dysplasia, and disorders associated with aberrant
 CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
 CC include degenerative disorders, growth deficiencies, hypoproliferative
 CC disorders, physical trauma, lesions, and wounds. An animal model
 CC preferably a mouse, in which a lats gene has been disrupted by homologous
 CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX

SO Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;

Query Match 25.8%; Score 812.6; DB 21; Length 3213;
 Best Local Similarity 65.1%; Pred. No. 1.le-200;
 Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTCACTGCCCGCCGCCAACACAGGCTACCGCGGTGACGCGCCGACACATC 1044
 DB 919 CACTCTGTAATCTTCAGCCTTCTGCCACTACAGTCACTGCATCACACCCGCTCCTATT 978
 QY 1045 CTTACACCTGTGAAGAGCGTGGTGTGCTGCGGCCGAGCCCGCCAGACAGCGCTGGGGCC 1104
 DB 979 CAACAGCCCGGTGAAGAGCATCGCGGTCTGAAACAGAGCTGCAGACTGCTTTAGCCCCA 1038
 QY 1105 TCGACACCCCGCTGGTGGCTGCGGCCACAGCACCTGCCACTGAGAGCGCTGGAGCAGAG 1164
 DB 1039 ACCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTTTCTGAG 1098
 QY 1165 GAGGCGACGCGAGCGCCACACCCCGCTGGATGTGGACTATGCGGCTCCGAGCGCAGGTGC 1224
 DB 1099 GGTACAGCTTCAAGTGTCCCTGTATCCACCTGTTGCTGAAGCTTCCAAGCTATCAAGGT 1158
 QY 1225 CCACCGCTCGTATCAAGACACTTGTCTGCTGCCAGTAGTCTGAGCAGTACAGCGTG 1284
 DB 1159 CCACCGCCGCTTATCCAAAACATCTGCTACACCAAAACCCATCTGCTCCCTCCATATAG 1218
 QY 1285 GACCTGGACAGCCTGTGCACCACTGTGCAGCAGAGTCTGCGAGGGGCGACTGTACAG 1344
 DB 1219 TCAGTAAGTAAAGCTGCAAGATGAACAGCTAGCTTACCAAGGAAGATGATAGTAG 1278
 QY 1345 GGGAGTGAAGAGCCACAAAGGTGCGAAGGAGAGCAAGCTGGCAGACAAAAGCAG 1404
 DB 1279 AAGAGTGGGACAGT-----GGTGACTCTGGGATATAAGAAAAGAAACAG 1323
 QY 1405 ATTACAGCTCCCGGTGCTTCCGGAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 DB 1324 ATTACAACTTCACTATCACTGTTCCGAAAACAAAGAGATGAAGAACGAAGAGTCT 1383
 QY 1465 CGCATCAAGAGTACTCCCTTATGCCCTTCAATTTCTTCAATGGAGCAACAGTGGAGAT 1524
 DB 1384 CGGATTAGAGTACTCCCAAGCGCTTTAAGTCTTCAATGGAGCAGCAGCTAGAGAAC 1443
 QY 1525 GTCATCAAAACCTACAGCAGAAAGTACGCGGAGGCTACAGCTGGAGCAGAAATGCC 1584
 DB 1444 GTCCTGAAGTCTCATCAGCAGCGTCTGCATCGGAAGAAGCAGCTAGAAAATGAATGATG 1503
 QY 1585 AAGCTGGGCTGTGTGAGCCGAGCAGGAGCAGATGAGGAGATCTCTACCAAGAGAG 1644
 DB 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAG 1563
 QY 1645 TCTAACTACAACCGGCTGAAGAGGCCAAGATGGACAGTCCATGTTTGTCAAAATCAAG 1704
 DB 1564 TCTAACTATATGCTTTAAAGGGCTTAAATGGACAGTCTATGTTTGTAAAGATAAG 1623
 QY 1705 ACTTAGGCATCGGTGCTTTGGGAAAGTGTGCTCGCTGTAAGCTGGACACTCACGCT 1764

DB 1624 ACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTAGCAAGAAAAGTGCATACAAAGCT 1683
 QY 1765 CTGTACGCCATGAGACTCTCAGAAAGAGAGATGCTTGAACCGGATCAAGTGGCCCAT 1824
 DB 1684 TTGTATGCAACAAGAGACTCTTCGAAAGAAAGACGCTTCTGCCGAAATCAGTGGCTCAT 1743
 QY 1825 GTCAGGCTGAGAGGAGCATCTCTGGCTGAAGCAGACAAATGAGTGGTGGTCTCAACTCTAC 1884
 DB 1744 GTGAAGCGGAGAGGATATCTTAGCAGAACCCACAAATGAGTGGTGGTCTCGCTGTAC 1803
 QY 1885 TACTCTTCCAGGAAAGAGCAGCTGTACTTGTGTGAGTACTATACACCGGCGGAT 1944
 DB 1804 TACTCTTCCAGGAAAGAGCAGCTGTACTTGTGTGAGTACTATACATCTCTGGGGGAT 1863
 QY 1945 ATGATGACCTGCTGATCAGATGAGTGTCTTCCCTGAGCAGCTGGCCCGTCTTACATTT 2004
 DB 1864 ATGATGACCTTATTAATAGAAATGGCATCTTCTGAAAATCTGGCAGATTTCTACATA 1923
 QY 2005 GCAGAGTTGACCTGGCCATTGAAAGTGTCCACAAGATGGCTTTATCCACCGGACATC 2064
 DB 1924 GCAGACTTACCTGTGCAAGTTGAAAGTGTCTAATAATGGCTTTATTCATAGATATTT 1983
 QY 2065 AAGCTGACAACTACTCATCGACTGGATGGTGTATTAAGCTGACAGATTTTGGCCTC 2124
 DB 1984 AAACCTGATAACATTTTGATTTGACCGTGTATGCCATATTAATAATGACTGACTTTGGCTG 2043
 QY 2125 TGCAGTGGATTCAGGTGAGCTCAAAATCCAAAGTACTACCAAGAAAGGAAACACATGAGA 2184
 DB 2044 TGCAGTGGCTTCAGATGAGCAGACTGACTCCAAAGTACTACCAAGTGGGGATGCCACGG 2103
 QY 2185 CAGGACAGCATGGAGCCCGGTGACTCTGGGACGATGTTTCCAACTGCTGCTGTGGAGAC 2244
 DB 2104 CAAGATAGCATGGATTTTCAGTAAAGATGGGAGATCTTCCATTTGCTGGTGTGGGAC 2163
 QY 2245 AGGTTAAAGACCCCTGGAGCAGAGGGCGCAGAACAGCAGCAGAGAGTGCCTGGCACAATTT 2304
 DB 2164 AGACTGAAGCCACTGGAGCGGAGAGTCTGCTGCCAGCAGCAGCGATGTCTAGCCCAATTT 2223
 QY 2305 CTTGTGGGACACCAAAATATACATCGCTCCGAGGTGCTTCCGCAAGAGGTACACGAG 2364
 DB 2224 CTGTTGGAGTCCCAATTTATTTGCACTGAAGTGTACTGCGAAGCAGGATATACACAG 2283
 QY 2365 CTCTGTGACTGGTGGAGCTGGTGTGATTTCTTTGAGATGCTGTTGGGAGCGGCT 2424
 DB 2284 CTGTTGACTGGTGGAGTGTGGTGTATTTCTTTGGAATGTTGGTGGGACACCTCT 2343
 QY 2425 TTTTGGCCCCCCCCACACAGCAGCAGCGAGCTGAAGGTGATCAACTGGGAGCAGCGCTG 2484
 DB 2344 TTTTGGCACAACCCCAATTTAGAAACACAAATGAAGGTTATCATCTGGCAAACTTCTCTA 2403
 QY 2485 CATATCCCTAGCAGGTGAGCTCAGCGCTCAGGCGCCGAGACCTCATCACGAAGCTGTGC 2544
 DB 2404 CATATCCCTCCTCAAGCTTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAAACTGTGT 2463
 QY 2545 TGGCGGCTGACTGCTGCGCTGGGAGGATGGGAGATGACCTCAAGCAGCAGCCGCTTC 2604
 DB 2464 CGAGCAGCAGAGACGCTCGGCAAGAACCGTGTGATGAGATAAAGGCTCATCCATTT 2523
 QY 2605 TTCAACACCATCTGACTTTTCCGTCAGATCCGAAAGCAGGCTGCACCTAGTCCCCCACC 2664
 DB 2524 TTTAAGACCATCGATTTCTCTAGTGTCTGAGACAGCAGTCTGCTTCTATCATATCCCTAAA 2583
 QY 2665 ATCAGCCACCCATGGACACCTCCAAATTTTGAACCGGTGGATGAAGAAAGCCCTGGCAG 2724
 DB 2584 ATCAGCTCAACAGATATACATCAATTTTCAGCCCTGTTGATCTGTGATAAATTTGGAGC 2643
 QY 2725 GAGCCAGCGAGAGAG---AGCGCCAAAGCCCTGGGACAGCTGGCTCCCCCAGCAGCAGAG 2781
 DB 2644 GATGGCAGGAGGAGGAAATATCAGTGACACTCTGAGCGGATGTTAATAAATTTGGAG 2703
 QY 2782 CATCCAGACGACGCTTCTATGATGAGTTCACCTTCGCGAGGTTCTTCGATGACAACGCTAT 2841

Fri Jan 17 11:17:03 2003

Db 2704 CACCCGAGCAGCCTTCTATGAGTTCACCTTTCGGAGGTTTTTGTGATGACAATGGCTAC 2763
 QY 2842 CCCTCCGGTGGCCGAGCCCTCAGAGCCGCGAGAGT 2880
 Db 2764 CCATATAATTATCAAGCCTATTGAGTATGATACATT 2802

RESULT 12
 AAT42119
 ID AAT42119 standard; cDNA; 3213 BP.
 XX
 AC AAT42119;
 DT
 XX
 XX 31-JAN-1997 (first entry)
 DE
 XX M-lats gene encoding large tumour suppressor.
 KW Mouse; m-lats gene; large tumour suppressor; fetal brain;
 KW protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX
 OS Mus musculus.
 XX
 XX Key Location/Qualifiers
 FH 1..2889
 CDS /tag- a
 FT /product= m-lats protein
 FT
 XX
 XX W09630402-A1.
 PN
 XX
 PD 03-OCT-1996.
 XX
 XX 26-MAR-1996; 96MO-US04101.
 PF
 XX
 PR 27-MAR-1995; 95US-0411111.
 XX
 XX (UYUA) UNIV YALE.
 PA
 PI Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX
 XX WPI: 1996-455275/45.
 DR P-PSDB; AAW05179.
 DR
 XX New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 PT
 XX Disclosure; Page 126-130; 215pp; English.
 PS
 XX This sequence encodes a mouse large tumour suppressor m-lats protein,
 CC and has been isolated from a newborn mouse brain phage lambda-ZAP
 CC cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
 CC (AAT42117). A homologous mouse sequence has also been isolated
 CC (m-lats2, AAT42120). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 XX Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;
 SQ

Query Match 25.78; Score 809.4; DB 17; Length 3213;
 Best Local Similarity 65.08; Pred. No. 7.6e-200;
 Matches 1235; Conservative 0; Mismatches 646; Indels 18; Gaps 2;
 QY 985 CAACCTGAGCCCTCACTCCCGCCGCCCAACACAGCTACCGCGGCGGCACACATC 1044

Db 919 CACTCTGCTAATCTCAGCCCTTCTGCCACTACAGTCACTGCGCATCACACCCGCTCCTATT 978
 QY 1045 CTTACCCCTTGAAGAGCGTGGTGTGCTGCGGCCGAGCCGAGAGCCGCTGGGGCCC 1104
 Db 979 CAACAGCCCGTGAAGAGCATGCGCGCTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCCA 1038
 QY 1105 TCGCACCCCGCTGGGTGGCTGCGCCACAGCAGCCTTGCACCTGAGAGCCTTGAGAGCAG 1164
 Db 1039 ACCATCTCTTGGATGCCACAGCCAGTTCAGACTGTTTCAGCCTACCCCTTTTCTGAG 1098
 QY 1165 GAGGCGAGCCGAGCCGACACCCGCTGGATGTGGATATGCGGCGCTCCGAGCGCAGGTGC 1224
 Db 1099 GGTACAGCTTCAAGTGTGCTGTCTATCCACCTGTTGCTGAAGCTCCAGCTATCAAGT 1158
 QY 1225 CCACCGCTCGTATCCAAAGCAGCTTCTGCTGCCAGTAAGTCTGAGAGTACAGCGTG 1284
 Db 1159 CCACCGCGCTTATCCAAACATCTGCTACACCAAAACCCATCTGCTCCCTCATATGAG 1218
 QY 1285 CACCTGAGCAGCCTGTGCACAGTGTGCAGAGTGTGGAGGGGCGACTGATCTAGAC 1344
 Db 1219 TCAGTAAGTAAGCCCTGCAAGATGAACGCTAGCTTACCCAAAGGAAGATGATAGT 1278
 QY 1345 GGGAGTGACAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGCGAGAGACAAAAAGCAG 1404
 Db 1279 AAGAGTGGGACAGT-----GGTACTCTGGGATAAAGAAAAAGAACAG 1323
 QY 1405 ATTACAGACTCCCGGTCCTGTCCGCAAGATAGCAGATGAGAGAGAGAGAGTCT 1464
 Db 1324 ATTACAACTTCACTTATCACTGTTTCCGAAACCAAGAAAGATGAAGAACAGAGAGTCT 1383
 QY 1465 CGCATCAAGATTACTCCGCTTATGCTTCAAAATTTCTTATGGAGCAACAGCTGGAGAAT 1524
 Db 1384 CGGATTGAGATTACTCCCGCAGAGCCCTTAAAGTTCTTATGGAGCAGCAGCTAGAGAAC 1443
 QY 1525 GTCATCAAAACCTACCCAGAGGTGAGCGGAGGCTACAGCTGGAGCAAGATGCCC 1584
 Db 1444 GTCTGAAGTCTCATCAGCAGCGTCTGCTCGAAGAGAGCAGCTAGAAAATGAATGATG 1503
 QY 1585 AAGCTGGCTCTGTGAGCCGAGCAGAGCAGATGAGAGAGATCCTCTACCAAGAGAG 1644
 Db 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAGAGATGCTTTGCCAGAGAG 1563
 QY 1645 TCTAACTACAACCGCTGAAGAGGCGCAAGATGGAGTCCATCTTTGTGAAAATCAAG 1704
 Db 1564 TCTAACTATATTCGTTTAAAGGGCTAAATGGACAGCTATGTTGTAAGATAAG 1623
 QY 1705 ACTTAGGCTCGTGGCTTGGGAGGTGCTGCTCGCTTGTAAAGCTGGACACTCACGCT 1764
 Db 1624 ACATTAGGAATAGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATATAAGCT 1683
 QY 1765 CTGTACCCCATGAAGACTCTCAGGAAGAGGATGTCCTGAACCGGANTCAAGTGCCCAT 1824
 Db 1584 TTGTATCAACAAAGACTTTCGAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCTCAT 1743
 QY 1825 GTCAAGCTGAGGGGACATCTTGGCTGAAGAGACAAATGAGTGGGTGGTCAAACTCTAC 1884
 Db 1744 GTGAAGCGGAGGGGATATCTAGCAGAGAGCCGACAAATGAGTGGGTGGTCCGCTGTAC 1803
 QY 1885 TACTCTTCCAGGACAAAGCAGCTCTACTTTGTGATGAGCTACATACACGCGGGAT 1944
 Db 1804 TACTCTTCCAGGACAAAGCAACTTGTACTTTGTGATGAGCTACATCTCTGGGGGGAT 1863
 QY 1945 ATGATGAGCTGTGATCAGGATGAGGTCTTCCCTCAGCAGCTGGCCGCTTCTACATT 2004
 Db 1864 ATGATGAGCTTATTAAATAGAAATGGGCACTTTCTCTGAAATCTGGCAGCATTTACATA 1923
 QY 2005 GCAGAGTTGACCCCTGGCCATTGAAAGTTCACAAGATGGGCTTTTATCCACCGGACATC 2064
 Db 1924 GCAGAACTTACCTGTGCAAGTGAAGTGTTCATAAAATGGGTTTTTATTCATAGAGATATT 1983
 QY 2065 AAGCCTGACACATCTACTCATCGACTGATGCTCATATTAGCTGACAGATTTTGGGCTC 2124
 Db 1984 AAACCTGATAACATTTTATTGATTCACCGTGTGATGCGCATATTAAATGACTGACTTTGGCTG 2043

Qy	2125	TGCACTGGATTCAGTGGAGTCAACAATTCAGTACTACCAAGAAAGGAAACACATGAGA	2184
Db	2044	TGCACTGGCTTCAGATGACACATCACTCAAGTACTACCAAGTGGGATCACCCACGG	2103
Qy	2185	CAGGACACATCGAGCCCGGTGACCTCTGGGACGATGTTTCCAACTGCTGCTGGAGAC	2244
Db	2104	CAAGATACATGGAATTCAGTAAACAAATGGGAGATCTTCCAAATGTCGTGGGGAC	2163
Qy	2245	AGGTTAAAGACCTGGAGCAGAGGGCGAGAGCAGAGAGTGGCTGGCAGCATCT	2304
Db	2164	AGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCAGCAGGATGCTAGCCCATCT	2223
Qy	2305	CTTGTGGGACACCAANTTACATCGCTCCGGAGTCTTCTCCGAAAGGTTACACGAG	2364
Db	2224	CTGGTTGGGACTCCCAATATATATTCACCTGAAGTCTACTGCGAAGAGGATATACACAG	2283
Qy	2365	CTCTGTGACTGGTGGAGCGTGGTGTGATTTCTTTGAGATGCTGGTGGGAGCGCT	2424
Db	2284	CTGTGTGACTGGTGGAGTGTGGTGTATTTCTTTGAAATGTTGGTGGGACAACTCT	2343
Qy	2425	TTCTTGGCCCCACCCACACAGAGAGCGAGCTGAAGTGTATCAACTGGGAGAGCAGCTG	2484
Db	2344	TTCTTGGCACAACCCCAATTTAGAACACAATGAAAGTTATCATCTGGCAAACTTCTTA	2403
Qy	2485	CATATCCCTAGCAGGTGAGGCTCAGCGCTGAGGCGCCGAGACCTCATCAGCAAGCTGTGC	2544
Db	2404	CACATCCCTCTCAAGCTAAGCTGAGTCTCTGAAGCTCTGACCTCATTTACAACTGTGT	2463
Qy	2545	TGCGGGGTGACTCGCGCTGGGAGGAGGATGGGAGATGACCTCAAGGCACACCGTTC	2604
Db	2464	CGAGGACACAGAGACCGCTCGGCAAGAACGCTGATGATGAGTAAAGGCTCATCAATT	2523
Qy	2605	TTCAACACCATCGATTTTCCGCTGACATCCGAAGCAGGCTGCACCTACGTCGCCACC	2664
Db	2524	TTTAAGACCATCGATTTCTTAGTGATCTGAGACAGAGCTGCTTCAATATCCCTAAA	2583
Qy	2665	ATCAGCCACCCCATGGACACCTCCCAATTTTGACCGGTGGATGAAGAACCCCTGGCAC	2724
Db	2584	ATCAGGCATCAACAGATACATCCCAATTTTCGACCTGTGATGATGAATTTGGGAGC	2643
Qy	2725	GAGGCCAGCGGAGAG-AGGCCCAAGCCCTGGGACAGCTGGCCCTCCCGCAGCAGCAAG	2781
Db	2644	GATGCGCAGGAGGAGAAATATCAGTGACACCTCTGAGCGGATGGTATATAAATGGGAG	2703
Qy	2782	CATCCAGACGAGCTTCTATGAGTTCACCTTCGCGAGGTTCTTCGATGACAAGCGCTAT	2841
Db	2704	CACCCGAGCAGCTTCTTATGAGTTCAGTTTCGGAGGTTTTTTGATGACAATGGCTAC	2763
Qy	2842	CCCTTCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAGT	2880
Db	2764	CCATATAATTATCCAAAGCCATTGAGTATGAATACATT	2802

RESULT 13

AAT42118

ID AAT42118 standard; cdna; 3984 BP.

XX AC AAT42118;

XX AC AAT42118;

XX DT 31-JAN-1997 (first entry)

XX DE H-lats gene encoding large tumour suppressor.

XX KW Human; h-lats gene; large tumour suppressor; fetal brain;

XX KW protein-serine/threonine-kinase; cell proliferation; antisense;

XX KW dominant-negative; cancer; degenerative disorder; trauma;

XX KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;

XX KW transgenic plant; transgenic animal; growth; senescence; ds.

XX OS Homo sapiens.

XX FH

Key Location/Qualifiers

FT CDS

231..3623

/*tag= a

/product= Lats protein

PN WO9630402-A1.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-0411111.

XX (UYUA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;

XX WPI; 1996-455275/45.

XX P-PSDB; AAW05178.

XX New isolated large tumour suppressor gene - used to develop prods.

XX for inhibiting cell proliferation or for enhancing proliferation

XX Claim 6; Page 118-123; 215pp; English.

XX This sequence encodes a human large tumour suppressor h-lats protein,

XX and has been isolated from a fetal human brain phage lambda-gt10

XX cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene

XX (AAT42119). The full-length cDNA is present in plasmid pBS(KS)-h-lats.

XX The gene encodes a putative protein-serine/threonine-kinase, and

XX inhibits cell proliferation and plays a crucial role throughout

XX development. Activators or inhibitors of lats function (e.g. an

XX antisense oligonucleotide or dominant-negative lats fragment) may be

XX used in therapy of cancer or other proliferative disorders,

XX degenerative disorders, trauma, growth deficiency, etc., and

XX fragments of the gene may be used as diagnostic probes. A

XX lats-inhibitor sequence may be expressed in a transgenic plant or

XX farm animal to confer increased growth and inhibit senescence.

XX SQ

Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match 24.1%; Score 760; DB 17; Length 3984;

Best Local Similarity 63.7%; Pred. No. 6e-187;

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

Qy 985 CAACCTGAGCCCTCACTCCCGCCCAACACAGCTGACGCGGTGACGCGCCGACACATC 1044

Db 1653 CACTCTGCTAATTCACGCTTCTGCTACAAAGTCACTGCAATTTACACAGCTCTTAT 1712

Qy 1045 CTTCAACCTGTGAAGAGCGTGTGCTGCGGCGCCGAGCCCGACAGCGGTGGGCCC 1104

Db 1713 CAACAGCCTGTGAAAGTATGCGTGTATTAACACAGAGCTACAGCTGCTTTAGCACCT 1772

Qy 1105 TCGCACCCCGCTGGTGGCTGGCCCGCCACAGCCTGCCACTGAGAGCCTGGAGACGAG 1164

Db 1773 ACACACCCCTTCTTGATACACAGCCCAATTCAAACTGTTCACCCAGTCTTTCTCTGAG 1832

Qy 1165 GAGGCGCAGCGCCCGCCACACCCCGCTGATGTGGACTATGCGGCTCCGAGCGCAGGTGC 1224

Db 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTGCTGAAGTCCAAATATCAAGGA 1892

Qy 1225 CCACCGCTTCGCTATCCAAAGCCTTGTGCTGCCAGTAAAGTCTGAGCAGTACAGCGTG 1284

Db 1893 CCACCAACCCCTACCCCAAAACATCTGCTGCACCAAAACCCATCTGTTCTTCCATACAG 1952

Qy 1285 GACCTGGACACCTGTGCACCGAGTGTGCAGCAGAGTCTGCGAGGGGCGCTGATCTAGAC 1344

Db 1953 TCAATCAGTAAGCCTAGCAAAAGAGATCAGCCAAAGTTCGCCAAGGAAGATGAGAGTGA 2012

Qy 1345 GGGAGTGACAGAGCCCAAGAGTTCGGAAGGAGGACAAAGCTGGCAGAGACAAAAGCAG 1404

Db 2013 AAGAGTTA-----TGAAATGTTGATAGTGGGGATAAAGAAAGAACAG 2057

Qy 1405 ATTCAGACCTCCCGGTCCTGTCGCGAAGAAATAGCAGAGATGAAGAGAGAGAGTCT 1464

Db 2058 ATTACAACTTACCTTACTTACTTAAAGAAACAAAGAGATGAAGAGCGAGGAATCT 2117
 QY 1465 CGCATCAAGAGTACTTCCCTTATCCCTTCAAAATCTTCATGGAGCAACACGCTGGAGAAT 1524
 Db 2118 CGTATTCAGAGTATTCCTCAGCATTTAAATCTTTATGGAGCAACATGTAGAAAT 2177
 QY 1525 GTCATCAAACTTACCAGCAGAGGTTCAGCCGGAGGTACAGCTGGAGCAGGAATGGCC 1584
 Db 2178 GTACTCAATCTCATCAGCAGCGCTTACATCGTAAACAAATAGAGAATGAATGATG 2237
 QY 1585 AAAGCTGGCTCTGTGAGGCGCAGCAGAGATGAGGAGATCTCTTACCAGAGGAG 1644
 Db 2238 CGGTTGATATCTCAGATGTCAGAGGTCCAGGATCAATAGAGAAGATGCTTGGCCAAAGAA 2297
 QY 1645 TCTAACTACAACCGGCTGAAGAGGCGCAAGATGAGAGATCTCTTACCAGAGGAG 1704
 Db 2298 TCTAATACATCCCTCTTAAAGGCTTAAATGAGCAAGTCTAATGTTGGAAGATAAG 2357
 QY 1705 ACTTAGGCATCGTGGCTTGGGAGTGTGCTCGCTTGAAGCTGGACACTCAGCT 1764
 Db 2358 ACATAGGAATAGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATAGGCT 2417
 QY 1765 CTGTACGCCATGAAGACTCTCAGAGAGAGGATGCTCTCAACCGGAATCAAGTGGCCAT 1824
 Db 2418 TTGTATGCAACAAACTCTTCAGAAAGAGTGTCTTCTTCCGAAATCAAGTCTCAT 2477
 QY 1825 GTCAGGCTGAGAGGACATCTTGGTGAAGCAGACATGAGTGGTGGTCAAACTCTAC 1884
 Db 2478 GTTAGGCTGAGAGATATCTTGGTGAAGCTGACATGAATGGTAGTCTATAT 2537
 QY 1885 TACTCTTCCAGGACAGCAGCTGTACTTGTGATGGACTACATACAGGCGGGAT 1944
 Db 2538 TATTCAATCCAGATGAAGACATTTATCTTGAATGGACTACATCTCTGGGGTAT 2597
 QY 1945 ATCATGAGCTGCTGATCAGGATGAGGTCTTCCCTGAGCAGCTGGCCGCTTCTACAT 2004
 Db 2598 ATGATGAGCTTAAATTAAGATGGCATCTTCCAGAAAGCTGGCAGATCTCATAT 2657
 QY 2005 GCAGAGTTGACCTGCCATTAAGTGTCCACAGATGGCTTTATCCCGGAGATC 2064
 Db 2658 GCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGTGTATTCATAGAGATAT 2717
 QY 2065 AGGCTGACACATCTCATCAGCTGGATGATTAAGCTGACAGATTTGGCCCT 2124
 Db 2718 AAACCTGATAATTTGATTGATCGTGGTGCATTAATTAAGTGAATCTTGGCCCT 2777
 QY 2125 TGCACTGATTCAGGTGAGCTCAAAATCCAACTACAGAAAGGAAACCATGAGA 2184
 Db 2778 TGCACTGCTTCAGATGACACAGATTTAAGTACTATCAGAGTGTGACCATCCAGG 2837
 QY 2185 CAGGACAGCTGAGGCGGCTGACCTCTGGGAGGATGTTTCCAACTGTCTGTGGAGAC 2244
 Db 2838 CAAGATAGCTGATTCAGTAAATGATGGGGATCCCTCAAGCTGTCTGTGGAGAC 2897
 QY 2245 AGTTAAAGACCTTGGAGCAGAGGGGAGAGAACAGCAGAGTGTCTGGCAGATCT 2304
 Db 2898 AGCTGAAGCCATAGAGCGGAGAGTGCACGCCAGCAGCAGTGTCTACCACTCT 2957
 QY 2305 CTGTGGGACACAAATACATCGCTCCGAGGCTTCTCCGCAAGAGGTTACAGCAG 2364
 Db 2958 TTGTTGGGACTCCCAATATATTCACCTGAAGTGTGCTACGAACAGGATACACAG 3017
 QY 2365 CTCTGTGCTGGTGGAGCTCGGTGTCTTTGAGATGCTGTTGGGAGCGGCT 2424
 Db 3018 TTGTGTGATTTGGAGTGTGGTGTATTTTGAATGTTGGTGGACACCTCT 3077
 QY 2425 TTCTTGGCCCCACCCACAGACGAGCTGAGGTTGATCACTGGGAGAGCAGCTG 2484
 Db 3078 TTCTTGGCAACACCATTAGAACAACAAATGAGGTTATCAACTGGCAACATCTCT 3137
 QY 2485 CATATCCCTACCGAGTGGCTCAGGCTGAGGCCCGGAGACCTTCATCAGCAAGCTGTC 2544

Db 3138 CACATTCACCACAACTAACTAGTCTCTGAAGCTTCTGATCTTATTAATACTTGC 3197
 QY 2545 TCGCGGCTGACTCGCGCTGGCAGGAGTGGGAGTACCTCAAGGACACACCGCTTC 2604
 Db 3198 CGAGGACCGGAGATCGCTTAGGCAAGATGGTCTGATGAATAAAGCTCATCCATTT 3257
 QY 2605 TTCAACACATCGACTTTTCCGCTGCATCCGATCCGATCCGATCCGATCCGATCC 2664
 Db 3258 TTAAACAATGACTTCTCCAGTGACCTGAGACAGAGTCTGCTTCATACATTCCTAAA 3317
 QY 2665 ATCAGCCACCCATGGACACTTCAATTTTACCCGCTGGATGAAGAAAGCCCTGGCAC 2724
 Db 3318 ATCAGACCCCAAGATACATCAATTTTATCTCTGATCTGATGAATAATTTGAT 3377
 QY 2725 GAGCCAGC---GGAGAGAGCGGAGGCTGGGACAGCTGGCCCTCCCGCAGCAGAA 2781
 Db 3378 GATGATACGAGGAGAAATGTAATGACACTCTCAATGGATGGTATAAAATGGAAG 3437
 QY 2782 CATCCAGCAGCAGCTTCTATGATGATTCACCTCCGAGGCTTCTTCATGACAGCGCTAT 2841
 Db 3438 CATCTCAACATCAATTTATGATTTACCTCCGAGGCTTTTGTGATGACAATGGCTAC 3497
 QY 2842 CCCTTCGGTGGCCGAGCC 2861
 Db 3498 CCATATAATTCGAGGCC 3517

RESULT 14

AAZ51505
 ID AAZ51505 standard; DNA; 3984 BP.
 AC AAZ51505;
 XX
 DT 21-JUN-2000 (first entry)
 XX Human Lats (large tumour suppressor) DNA.

Human Lats; large tumour suppressor; cytostatic; vulnary;
 cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
 treatment; prevention; screening; cancer; skin; ovarian tumour;
 soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
 LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
 dysplasia; degenerative disorder; growth deficiency; physical trauma;
 hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.
 OS Homo sapiens.

Key Location/Qualifiers
 CDS 231..3623
 FT /*tag= a
 FT /product= "Lats protein"

WO200010602-A1.

02-MAR-2000.

18-AUG-1999; 99WO-US19068.

18-AUG-1998; 98US-0096996.

18-AUG-1998; 98US-0096997.

(UYU) UNIV YALE.

Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

Turechalk GS, Stewart RA;

WPI: 2000-246496/21.

P-PSDB; AAY70390.

Use of lats proteins, complexes of lats and cdc2 for treating cancer

that is refractory to treatment by standard chemotherapy and radiation
 therapy, and disorders associated with aberrant levels of cdc2 activity

120-121: 137pp: English:

Db 2402 CCGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAGAA 2461

QY 1645 TCTAACTACACCGGCTGAAGAGGCCAAGATGGACAAAGTCCATGTTGTGAAAATCAAG 1704

Db 2462 TCTAATTACATCCGCTTAAAGGGCTAAATGGACAGTCTATGTTGTGAGATAAAG 2521

QY 1705 ACTCTAGGACGGTGGCTTTGGGGAAGTGTCCCTGTGTAAGCTGGACACTCACGGT 1764

Db 2522 ACACCTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACTAAGGCT 2581

QY 1765 CTGTAGCCCATGAACACTCTCAGGAAGAGGATGCTTGAACCCGAATCAAGTGCCCAT 1824

Db 2582 TTGTATGCAACAAAACCTCTTCAAGAAAGATGTTCTTTCGAAATCAAGTCGCTCAT 2641

QY 1825 GTCAAGGCTGAGAGGACATCTCGCTGAAGCAGACAAATAGTGGGTGGTCAAACTCTAC 1884

Db 2642 GTTAAGGCTGAGAGATATCTCGCTGAAGCTGACATGAATGGGTAGTTCGTCTATAT 2701

QY 1885 TACTCTTCCAGGACAAGACAGCCTGTACTTTGTGATGGACTACATACCAGCGGGGAT 1944

Db 2702 TATTCAATCCAAAGATAAGACAAATTTATCTTTGTAATGGACTACATCTCTGGGGTGTAT 2761

QY 1945 ATGATGAGCCTCTCATCAGGATGAGGTCTTCCCTGAGCACCTGGCCCGCTTCTACATT 2004

Db 2762 ATGATGAGCCTATTAAATAGAAATGGCATCTTTCAGAAAGTCTGGCACGATTTCTACATA 2821

QY 2005 GCAGAGTTGACCTGGCCTGAAAGTGTCCACAAAGATGGGCTTTATCCACGGGACATC 2064

Db 2822 GCAGAACTTACCTGTGCAAGTGAAGAGTGTTCATAAATGGGTTTATTTCATAGAGATT 2881

QY 2065 AAGCCTGACAACTACTCATCGACCTGGATGGTCTATATTAAGCTGACAGATTTTGGCCCTC 2124

Db 2882 AAACCTGATAATATTTTGATGATCGTGTGCTATATTAATTAAGTGAAGTCTGGCCCTC 2941

QY 2125 TGCACCTGATTCAGGTGACACTACAAATTCCTAAGTACTACCAGAAAGGAACACATGAGA 2184

Db 2942 TGCACCTGCTTCAGATGGACACACAGATTTCTAAGTACTATCAGAGTGGTGACCATCCAGG 3001

QY 2185 CAGGACAGCATGGAGCCGGTGACCTCTGGGACGATGTTTCCAACTGTGCTGGGAGAC 2244

Db 3002 CAAGATAGCAGTGGATTCAGTAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 3061

QY 2245 AGTTTAAAGCCTGGAGAGCGGCGAGAGACAGCAGAGGTGCTGGCAGACATTCT 2304

Db 3062 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCACAGCAGCGATGTCTAGCACATTCT 3121

QY 2305 CTGTGCGGACACCAAAATTAATCGCTCCGAGGTGCTTCCCGCAAGAGGTACAGCAG 2364

Db 3122 TTGGTTGGGACTCCCAATTAATATGCACTGAAGTGTGCTACGAAAGAGATACACACAG 3181

QY 2365 CTCGTGACTGTGTGAGCGTGGTGTGATCTCTTTGATGATGCTGGTTGGCAGCGCT 2424

Db 3182 TTGTGTGATTTGGTGGAGTGTGGTGTATCTTTTGAATGTTGGTGGGACCACTCT 3241

QY 2425 TTCCTGGCCCCACCCACAGAGCGAGCTGAAGTGTATCAACTGGGAGAGACAGCTG 2484

Db 3242 TTCCTGGCACAACACCACTTAGAAACAAATGAAGTTATCAACTGGCAAACTCTCT 3301

QY 2485 CATATCCCTAGCAGGTGAGGCTAGCGCTGAGGCCCGAGACCTCATCAGGAAGCTGTGC 2544

Db 3302 CACATTCACCAACAGCTAAACTCAGTCTGAAGCTTCTGATCTTATTATAAACTTTGC 3361

QY 2545 TGCGGGCTGACTCGGCTGGGAGGATGGGAGATGACCTCAAGGCACACCCGTTTC 2604

Db 3362 CGAGACCCGAGATCGGTTAGGAAGAAATGGTGTGATGAATAAAGCTCATCTATT 3421

QY 2605 TTCACACCATCGACTTTTCCGTCGACATCCGAAAGAGGCTGCACCTACGTCGCCACC 2664

Db 3422 TTTAAACAATTGACTTCTCCAGTGACCTGAGACAGAGCTGTGCTCATACATTCCTAAA 3481

QY 2665 ATCAGCCACCCATGGACACCTTCAATTTTACCGCGGTGGATGAAGAAAGCCCTGGCAC 2724

Db 3482 ATCACACACCAACAGATACATCAAAATTTTGTATCTGTTGATCCTGTGATAAATATGGAGT 3541

QY 2725 GAGGCCAGC---GGAGAGAGCGCCAAAGCGCTGGGACACGCTGGCCTCCCCAGCAGCAAG 2781

Db 3542 GATGATAACGAGGAGAAAATGTAATGACACTCTCAATGGATGATATAAAAATGGAAG 3601

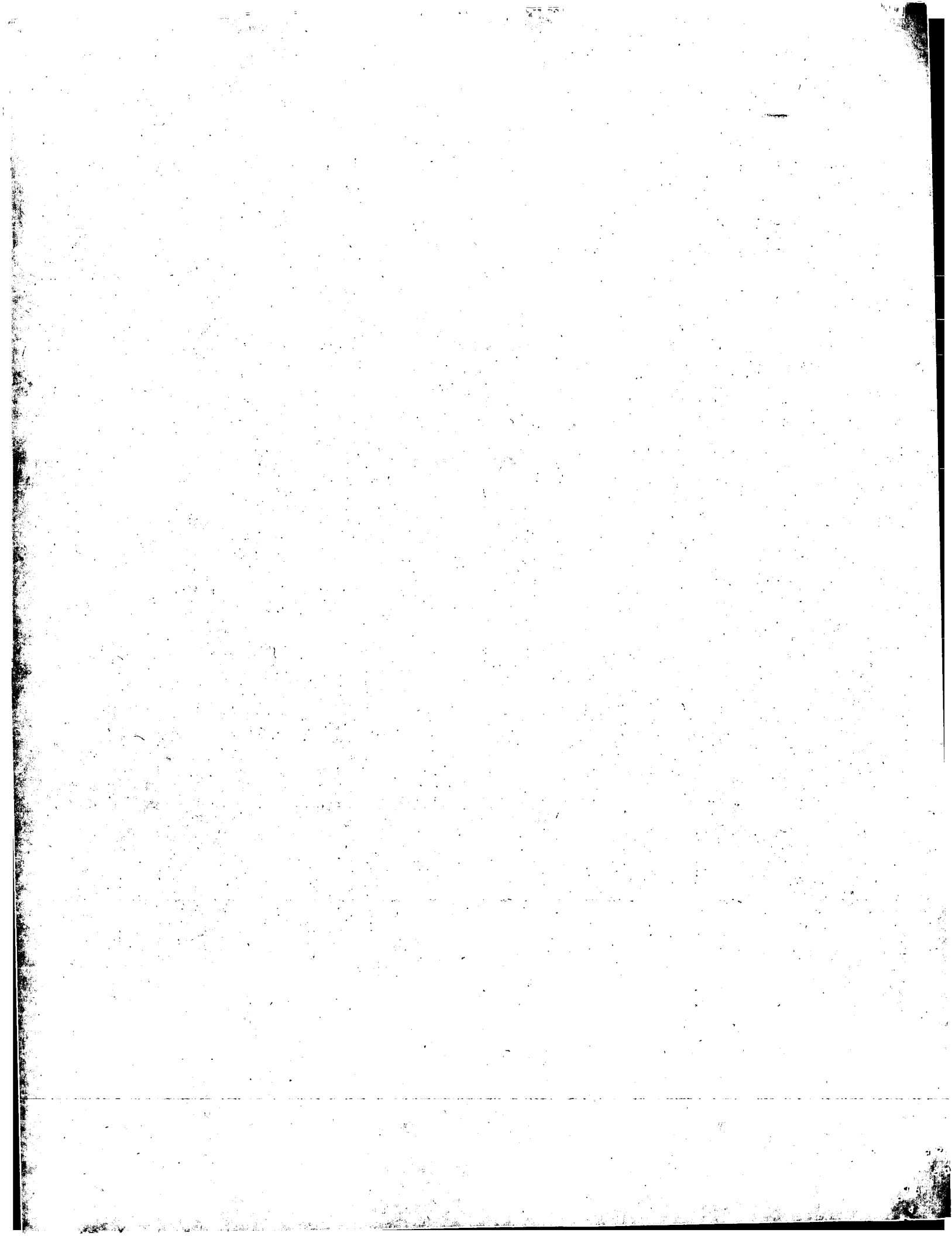
QY 2782 CATCCAGAGCAGCGCTTCTATGATGTTTCCAGGTTCTTCGATGACAAACGGCTAT 2841

Db 3602 CATCTCTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTGATGACAATGGCTAC 3661

QY 2842 CCCTTCCGGTCCCCGAAGCC 2861

Db 3662 CCATATAATTATCCGAAGCC 3681

Search completed: January 16, 2003, 10:34:06
Job time : 464.423 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:02 ; Search time 64.1915 seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagaccacccgaagtt.....aagagcattatttggggg 3155

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3155	100.0	3155	4 US-09-442-100-7	Sequence 7, Appli
2	1293	41.0	1961	4 US-09-509-902A-15	Sequence 15, Appli
3	1148.4	36.4	1498	4 US-09-509-902A-6	Sequence 6, Appli
4	812.6	25.8	3213	4 US-09-442-100-5	Sequence 5, Appli
5	760	24.1	3984	4 US-09-442-100-3	Sequence 3, Appli
6	564.6	17.9	5720	4 US-09-442-100-1	Sequence 1, Appli
7	333	10.6	638	4 US-09-328-111-26	Sequence 26, Appli
8	258.8	8.2	2160	4 US-09-588-256-1	Sequence 1, Appli
9	233.8	7.4	1935	2 US-08-878-989-11	Sequence 11, Appli
10	233.8	7.4	1935	2 US-09-272-796-11	Sequence 11, Appli
11	225.2	7.1	3018	2 US-08-860-150-6	Sequence 6, Appli
12	225.2	7.1	3018	2 US-08-338-132-6	Sequence 6, Appli
13	222.6	7.1	2101	2 US-08-860-150-1	Sequence 1, Appli
14	222.6	7.1	2101	3 US-09-338-132-1	Sequence 1, Appli
15	214.4	6.8	678	4 US-09-328-111-66	Sequence 1, Appli
16	161.8	5.1	2549	4 US-09-467-082-3	Sequence 66, Appli
17	149.8	4.7	2726	2 US-08-422-699A-12	Sequence 3, Appli
18	149.8	4.7	2726	2 US-08-422-706B-12	Sequence 12, Appli
19	144.4	4.6	3182	1 US-08-484-044-11	Sequence 12, Appli
20	143.6	4.6	1599	3 US-09-256-465-1	Sequence 1, Appli
21	143.6	4.6	1599	4 US-09-167-322-3	Sequence 3, Appli
22	143.4	4.5	2511	2 US-08-422-699A-8	Sequence 8, Appli
23	143.4	4.5	2511	2 US-08-422-706B-8	Sequence 8, Appli
24	136.4	4.0	2610	2 US-09-212-771-1	Sequence 1, Appli
25	126.4	4.0	2610	3 US-09-091-058-1	Sequence 1, Appli
26	123.8	3.9	4739	3 US-08-685-871-1	Sequence 1, Appli
27	121.8	3.9	5053	2 US-08-685-576-2	Sequence 2, Appli

Sequence 5, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 2, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 10, Appli
Sequence 1, Appli
Sequence 10, Appli
Sequence 10, Appli
Patent No. 5266464

ALIGNMENTS

RESULT 1
US-09-442-100-7
Sequence 7, Application US/09442100
Patent No. 6359193
GENERAL INFORMATION:
APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Weiyl
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2943

Sequence 7, Appli
Sequence 15, Appli
Sequence 6, Appli
Sequence 5, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 26, Appli
Sequence 1, Appli
Sequence 11, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 1, Appli
Sequence 66, Appli
Sequence 3, Appli
Sequence 12, Appli
Sequence 12, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 2, Appli

US-09-442-100-7

Query Match		100.0%	Score 3155;	DB 4;	Length 3155;		
Best Local Similarity		100.0%	Pred. No. 0;				
Matches 3155; Conservative 0;		Mismatches 0;	Indels 0;	Gaps 0;			
QY	1	ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATTCGATATTC	60				
Db	1	ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATTCGATATTC	60				
QY	61	CTCCTGCCCTTTGCCAACAGTTCAGCACTTCGGCAGCTGCGAGAGTGAACCGGCAGATG	120				
Db	61	CTCCTGCCCTTTGCCAACAGTTCAGCACTTCGGCAGCTGCGAGAGTGAACCGGCAGATG	120				
QY	121	CTTCAGGAGTTGGTGAATGCGGCATGTGACAGAGAGTGGCTGGCAGAGCGCTCACGCAG	180				
Db	121	CTTCAGGAGTTGGTGAATGCGGCATGTGACAGAGAGTGGCTGGCAGAGCGCTCACGCAG	180				
QY	181	ACGGCAGTAGGAGTATCGAAGCTGCCTTGGAGTACATCAGTAAGATGGGTACCTGGAC	240				
Db	181	ACGGCAGTAGGAGTATCGAAGCTGCCTTGGAGTACATCAGTAAGATGGGTACCTGGAC	240				
QY	241	CCCAGGAATGAGCAGAGATTGCGAGTATCAAGCAGACCTCCCCAGGAAGGCGCTGGCG	300				
Db	241	CCCAGGAATGAGCAGAGATTGCGAGTATCAAGCAGACCTCCCCAGGAAGGCGCTGGCG	300				
QY	301	TCCACCCCGGTGACTGCGGGCCAGTTTCAGGGCAGCAGGGGAAGCACTCCCATCTCTAC	360				
Db	301	TCCACCCCGGTGACTGCGGGCCAGTTTCAGGGCAGCAGGGGAAGCACTCCCATCTCTAC	360				
QY	361	CACAGCTGGGTGGTGCAAACTACGAGGGCCCCCGCCACCTGGAGGAGATGCGCGGGCAA	420				
Db	361	CACAGCTGGGTGGTGCAAACTACGAGGGCCCCCGCCACCTGGAGGAGATGCGCGGGCAA	420				
QY	421	TATTTAGACTTTCTCTCCCTGGAGCCGAGCCGACCCACCGTGCAGAGGTCCACGAG	480				
Db	421	TATTTAGACTTTCTCTCCCTGGAGCCGAGCCGACCCACCGTGCAGAGGTCCACGAG	480				
QY	481	CATCTCTCCCAAGGGTTACAGCAGCAGTAGAGCCAAAGTCGCGACATTTCCGGGACACAC	540				
Db	481	CATCTCTCCCAAGGGTTACAGCAGCAGTAGAGCCAAAGTCGCGACATTTCCGGGACACAC	540				
QY	541	TATGGTCTGTGGTCACTACTATCGGAGCAGTCTGGGTATGGGTGCGAGCGCTTCCTCC	600				
Db	541	TATGGTCTGTGGTCACTACTATCGGAGCAGTCTGGGTATGGGTGCGAGCGCTTCCTCC	600				
QY	601	TTCCAGAACAGAGCCACAGATGCTTATTCAGCAGTGCACAGGCGCCAGGGTGGCCCT	660				
Db	601	TTCCAGAACAGAGCCACAGATGCTTATTCAGCAGTGCACAGGCGCCAGGGTGGCCCT	660				
QY	661	CCGCGCAGGCTCACCCTTTCTGGCCATGCTGGGCTGTACACTGCTCGCACCAAGCGG	720				
Db	661	CCGCGCAGGCTCACCCTTTCTGGCCATGCTGGGCTGTACACTGCTCGCACCAAGCGG	720				
QY	721	GGGGCTACCCCACTTGGGGCCACCATATCATGTGTGGCACCCTGGGGTCCCAAGCTTT	780				
Db	721	GGGGCTACCCCACTTGGGGCCACCATATCATGTGTGGCACCCTGGGGTCCCAAGCTTT	780				
QY	781	ACTGGGAAAGCTCTGCACAGGCTGTGCTGGCACCGCTCCAGGAACAGGCTCAATGCTGAC	840				
Db	781	ACTGGGAAAGCTCTGCACAGGCTGTGCTGGCACCGCTCCAGGAACAGGCTCAATGCTGAC	840				
QY	841	TTGTACGAGCTGGGCTCCAGGTCGCCCTGGTCTGCGAGTCTCACTGGCAGCGCGGACTCG	900				
Db	841	TTGTACGAGCTGGGCTCCAGGTCGCCCTGGTCTGCGAGTCTCACTGGCAGCGCGGACTCG	900				
QY	901	CTCCAGAAAGGCTTAGAAGCCTCGCGGCCGATGTGGCTTTTCGGGCTGGCCCGAGC	960				
Db	901	CTCCAGAAAGGCTTAGAAGCCTCGCGGCCGATGTGGCTTTTCGGGCTGGCCCGAGC	960				
QY	961	AGGACCAACTCTTCAACAAACCAACTGAGCCCTCACTGCCGCGCCCGCAACAGGTC	1020				
Db	961	AGGACCAACTCTTCAACAAACCAACTGAGCCCTCACTGCCGCGCCCGCAACAGGTC	1020				

QY	1021	ACGGCCTGACGGCGGCACACATCTTTCACCTGTGAAGAGCGTGGTGTCTCGGGCC	1080
DB	1021	ACGGCCTGACGGCGGCACACATCTTTCACCTGTGAAGAGCGTGGTGTCTCGGGCC	1080
QY	1081	GAGCCCCAGACAGCGGTGGGGCCCTCGCACCCCGCTGGGTGGTGGTGGTGGTGGT	1140
DB	1081	GAGCCCCAGACAGCGGTGGGGCCCTCGCACCCCGCTGGGTGGTGGTGGTGGTGGT	1140
QY	1141	GCCACTGAGAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1200
DB	1141	GCCACTGAGAGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1200
QY	1201	TATGGCGCTCCGAGCGAGTGGCCCGCTCGCACCCCGCTGGGTGGTGGTGGTGGT	1260
DB	1201	TATGGCGCTCCGAGCGAGTGGCCCGCTCGCACCCCGCTGGGTGGTGGTGGTGGT	1260
QY	1261	AGTAAGTCTGAGCAGTACAGCTGGAGCTGACAGCGCTGTCACACAGTGTGACAGAG	1320
DB	1261	AGTAAGTCTGAGCAGTACAGCTGGAGCTGTCACAGCTGTGACACAGTGTGACAGAG	1320
QY	1321	CTGGAGGGGCACTGATCTAGACGGGAGTGACAAGAGGAGGAGGAGGAGGAGGAG	1380
DB	1321	CTGGAGGGGCACTGATCTAGACGGGAGTGACAAGAGGAGGAGGAGGAGGAGGAG	1380
QY	1381	AAAGCTGGCAGAGCAAAAAGCAGATTTCAGACCTCCCGCTGGTGGTGGTGGTGGT	1440
DB	1381	AAAGCTGGCAGAGCAAAAAGCAGATTTCAGACCTCCCGCTGGTGGTGGTGGTGGT	1440
QY	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTTCCCTTTTCAAAATTC	1500
DB	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTTCCCTTTTCAAAATTC	1500
QY	1501	TTTCATGAGCAACACCTGGAGATGTCATCAAACTTACAGAGAGGAGTACAGCGAGG	1560
DB	1501	TTTCATGAGCAACACCTGGAGATGTCATCAAACTTACAGAGAGGAGTACAGCGAGG	1560
QY	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGCGCGAGCAGAGAGATG	1620
DB	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGCGCGAGCAGAGAGATG	1620
QY	1621	AGGAAGCTCTTACAGAGGAGTCTAATAACACCGCTGAAGAGGAGGAGGAGGAGG	1680
DB	1621	AGGAAGCTCTTACAGAGGAGTCTAATAACACCGCTGAAGAGGAGGAGGAGGAGG	1680
QY	1681	AAGTCCATGTTGTGAAATCAAGACTCTAGCATCGGCTTTGGGAAAGTGTGCCTC	1740
DB	1681	AAGTCCATGTTGTGAAATCAAGACTCTAGCATCGGCTTTGGGAAAGTGTGCCTC	1740
QY	1741	GCTTGTAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGTC	1800
DB	1741	GCTTGTAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGTC	1800
QY	1801	CTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGATCCTGGCTGAAGCAGAC	1860
DB	1801	CTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGATCCTGGCTGAAGCAGAC	1860
QY	1861	AATGAGTGGGTGGTCAAACTCTACTCTCTTCCAGGACAAAGGAGCAGCTGTACTTTG	1920
DB	1861	AATGAGTGGGTGGTCAAACTCTACTCTCTTCCAGGACAAAGGAGCAGCTGTACTTTG	1920
QY	1921	ATGGACTACATACAGCGGGGATATGATGAGCTGTGATCAGGATGAGGCTTCTCCCT	1980
DB	1921	ATGGACTACATACAGCGGGGATATGATGAGCTGTGATCAGGATGAGGCTTCTCCCT	1980
QY	1981	GAGCAGCTGGCGCTTCTACATTCAGAGTGTACCTGGCCATGAAAGTGTCCACAG	2040
DB	1981	GAGCAGCTGGCGCTTCTACATTCAGAGTGTACCTGGCCATGAAAGTGTCCACAG	2040
QY	2041	ATGGCTTTTATCCACCGGAGATCAAGCTGACAACTACTCATCGACCTGGATGGTCAT	2100
DB	2041	ATGGCTTTTATCCACCGGAGATCAAGCTGACAACTACTCATCGACCTGGATGGTCAT	2100

```

QY 2101 ATTAAGTACAGAGATTTGGCTCTCCACTGATTCAGGTGAGTCCACAAATCCAAAGTAC 2160
Db 2101 ATTAAGTACAGAGATTTGGCTCTCCACTGATTCAGGTGAGTCCACAAATCCAAAGTAC 2160
QY 2161 TACCAGAAAGGACACATGAGACAGACAGCATGAGTGGTCCGCTCTGGGACGAT 2220
Db 2161 TACCAGAAAGGACACATGAGACAGACAGCATGAGTGGTCCGCTCTGGGACGAT 2220
QY 2221 GTTTCACACTGCTGCTGTGAGACAGGTTAAAGACCTCTGAGCAGAGGCGGAGAGCAG 2280
Db 2221 GTTTCACACTGCTGCTGTGAGACAGGTTAAAGACCTCTGAGCAGAGGCGGAGAGCAG 2280
QY 2281 CACGAGAGTGGCTGACACATTTCTTTGTCGGGACACCAAAATACATCCTCCGGAGGTG 2340
Db 2281 CACGAGAGTGGCTGACACATTTCTTTGTCGGGACACCAAAATACATCCTCCGGAGGTG 2340
QY 2341 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGAGGCTGCGGTGTGATTTCTTTT 2400
Db 2341 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGAGGCTGCGGTGTGATTTCTTTT 2400
QY 2401 GAGATGCTGTGGGACGCGCTTTCTTTGGCCCGCCACCCACACAGAGCAGCTGAAG 2460
Db 2401 GAGATGCTGTGGGACGCGCTTTCTTTGGCCCGCCACCCACACAGAGCAGCTGAAG 2460
QY 2461 GTGATCAACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
Db 2461 GTGATCAACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
QY 2521 CGAGACCTCATCAGNAAGTGTGCTGCGCGGTGACTGCGCGCTGGGAGGAGTGGGCA 2580
Db 2521 CGAGACCTCATCAGNAAGTGTGCTGCGCGGTGACTGCGCGCTGGGAGGAGTGGGCA 2580
QY 2581 GATGACCTCAAGGACACACCGCTTTCTTCAACACCATCTGACTTTTCCGTCACATCCGAAG 2640
Db 2581 GATGACCTCAAGGACACACCGCTTTCTTCAACACCATCTGACTTTTCCGTCACATCCGAAG 2640
QY 2641 CAGGCTGACCTTACGCTCCCGCCACCATCAGCCACCCCATGAGACCTTCCAAATTTGACCGG 2700
Db 2641 CAGGCTGACCTTACGCTCCCGCCACCATCAGCCACCCCATGAGACCTTCCAAATTTGACCGG 2700
QY 2701 GTGGATGAAGAAGCCCTTGGACAGGCGCCAGGAGAGAGAGCCCTGGGACAGC 2760
Db 2701 GTGGATGAAGAAGCCCTTGGACAGGCGCCAGGAGAGAGAGCCCTGGGACAGC 2760
QY 2761 CTGGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGATTCACCTTCCGACGG 2820
Db 2761 CTGGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGATTCACCTTCCGACGG 2820
QY 2821 TTCTTCGATGACAACGGCTATCCCTTCCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAT 2880
Db 2821 TTCTTCGATGACAACGGCTATCCCTTCCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAT 2880
QY 2881 GCAGACCCAGGGATGCGGACTTGAAGTGGCGGAGGCTGCCAGCCGCTGTACGTG 2940
Db 2881 GCAGACCCAGGGATGCGGACTTGAAGTGGCGGAGGCTGCCAGCCGCTGTACGTG 2940
QY 2941 TAAGCTCAGTTAAACACAACTCGAGGAACCCAAATGAGATTTCTTTTCAGAGACAA 3000
Db 2941 TAAGCTCAGTTAAACACAACTCGAGGAACCCAAATGAGATTTCTTTTCAGAGACAA 3000
QY 3001 ACTCAAGCTTAGGAATCCCTTCAATTTTATGTTCTGTTAAATGGGCAACGAGAGTCAAC 3060
Db 3001 ACTCAAGCTTAGGAATCCCTTCAATTTTATGTTCTGTTAAATGGGCAACGAGAGTCAAC 3060
QY 3061 ATGATTTCAATTAGCCCTCTGAGACCTTCACTGATTAACACAGTATTTTAAAAA 3120
Db 3061 ATGATTTCAATTAGCCCTCTGAGACCTTCACTGATTAACACAGTATTTTAAAAA 3120
QY 3121 TTAGTACAGTATGGAAGAGCACTTATTTTGGGG 3155
Db 3121 TTAGTACAGTATGGAAGAGCACTTATTTTGGGG 3155

```

```

RESULT 2
US-09-509-902A-15
; Sequence 15, Application US/09509902A
; Patent No.: 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1961
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-15

```

```

Query Match 41.0%; Score 1293; DB 4; Length 1961;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 255; Indels 21; Gaps 3;

```

```

QY 1185 CCCGCTGGATGTGACTATGCGGCTCCGAGCGCAGGTGCCACCGCTCCCGTATCCAAA 1244
Db 2 CCCGCTGGACGTGGAGTACGGAGGCCACGCGGAGGTCCCGCTCCCGCTACCCGAA 61
QY 1245 GCACCTGTGCTGCCAGTAACTCTGAGCAGTACAGCGTGGACCTGGAGCAGCTGTGCAC 1304
Db 62 GCACCTGTGCTGCCAGCAAGTCGAGCAGTAC-----GACCTGGACAGCTGTGCGC 115
QY 1305 CAGTGTGACGAGAGTCTGCGAGGGG---CACTGATCTAGAGCGGAGTGAACAAGCCA 1361
Db 116 AGGCATGGAGCAGAGCTCCGTCGCGGCCCAACAGCGCGGCGGCGCAAGAGCGG 175
QY 1362 CAAAGTGGGAAGGAGCAAAAGCTGGCAGACACAAAAGAGATTCAGACCTCCCGGT 1421
Db 176 CAAAGCGCAAGGGGACAAAGCGGGAAGATTAAGAAGAGATTCAGACCTTCCTCGT 235
QY 1422 GCCTTCGCGCAAGAATAGCAGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTACTC 1481
Db 236 TCCGCTCCGCAAAAACAGCAGAGAGAGAGAGAGAGTCAAGCATCAAGAGTACTC 295
QY 1482 CCCTTATGCTTCAATTTCTTCATGAGCAACAGCTGGAGATGTCATCAAAACCTACCA 1541
Db 296 GCATACGCTTTAAGTTCTTCATGGAGCAGCAGTGGAGATGTCATCAAAACCTACCA 355
QY 1542 GCAGAGGTTCAGCGGAGGCTACAGTGGAGCAGGAAATGGCAAGCTGGGCTCTGTGA 1601
Db 356 GCAGAGGTTCAGCGGAGGCTACAGTGGAGCAGGAAATGGCAAGCTGGGCTCTGTGA 415
QY 1602 GCGCAGCAGGAGCAGATGAGGAGATCCTTACAGAGAGAGTCTAACTACAACCGGT 1661
Db 416 AGCTGAGCAGGAGCAGATGCGGAAGATCTCTTACAGAAAAGAGTCTAATTACAACAGGT 475
QY 1662 GAAGAGGCCAAGATGGACAAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGC 1721
Db 476 AAAGAGGCCAAGATGGACAAGTCTATGTTGTCAAGATCAAAACCTGGGATCGGTGC 535
QY 1722 CTTTGGGGAAGTGTGCTCGCTTGTAAAGCTGGACACTCACGCTCTGTACGCCATCAAGAC 1781
Db 536 CTTTGGAGAGTGTGCTTGTGTTAGGTGACACTCACGCCCTGTACGCCATCAAGAC 595
QY 1782 TCTCAGGAAGAGAGTGTCTGAACCGGAATCAAGTGGCCCCCATGTCAAGGCTGAGAGGA 1841
Db 596 CTTAAGAAAAAGAGTGTCTGAACCGGAATCAAGTGGCCCCCATGTCAAGGCTGAGAGGA 655
QY 1842 CATCTGCTGAGCAGACATGAGTGGTGGTCTAACTTACTTACTTCTTCCAGAGCAA 1901
Db 656 CATCTGCGCGAGGAGACAATGAGTGGTGGTCTAACTTACTTACTTCTTCCAGAGCAA 1951

```

```

QY 1902 GGAACAGCCTGTACTTTGTGTGATGAGTACATACACAGGGGGGATATGATGAGCCTGCTGAT 1961
Db 716 AGACAGCCTGTACTTTGTGTGATGAGTACATACATCCCTGGTGGGACATGATGAGCCTGCTGAT 775
QY 1962 CAGGATGAGGCTTCTCCCTGAGCAGCCTGGCCCGCTTCTACATTCGACAGTGTACCCCTGGC 2021
Db 776 CCGGATGAGGCTTCTCCCTGAGCAGCCTGGCCCGCTTCTACATTCGACAGTGTACCCCTGGC 835
QY 2022 CATTGAAAGTGTCCACAGATGGGCTTTATCCACCGGGGACATCAAGCCTCAGACATACCT 2081
Db 836 CATTGAGAGTGTCCACAGATGGGCTTTATCCACCGGGGACATCAAGCCTCAGACATACCT 895
QY 2082 CATGACCTGTGATGCTATATTAAGCTGACAGATTTTGGCCTGTGACCTGATTCAGGTG 2141
Db 896 GATGATGCTGATGCTGATTAAGCTGACAGATTTTGGCCTGTGACCTGATTCAGGTG 955
QY 2142 GACTCACAATTTCCAAAGTACTACCAAGAGGAGGACACATGAGACAGACAGATGAGCC 2201
Db 956 GACTCACAATTTCCAAATATTACCAAGAGGAGGACATGAGACAGACAGATGAGCC 1015
QY 2202 CGGTGACCTCTGGGACGATGTTTCCAACTGCTGTGGAGACAGGTTAAAGACCTTGA 2261
Db 1016 CAGGACCTCTGGGATGATGCTTAACTGCTGTGGAGACAGGCTCAAGACCTTGA 1075
QY 2262 GCAGAGGCGGCAAGCAGCAGCAGGCTGCTGAGCAGATTTCTTGTGCGGACACCAA 2321
Db 1076 GCAGAGGCGGCAAGCAGCAGGCTGCTGAGCAGATTTCTTGTGCGGACACCAA 1135
QY 2322 TTACATGCTGCTGGAGGCTGCTTCTCGCAAGGAGTACAGCAGCTGCTGAGCTGTGAG 2381
Db 1136 CTACATGCTGCTGGAGGCTGCTTCTCGCAAGGAGTACAGCAGCTGCTGAGCTGTGAG 1195
QY 2382 CGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2441
Db 1196 TGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
QY 2442 CACAGAGCAGCTGAGGCTGATCAACTGGGAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2501
Db 1256 CACAGAGCAGCTGAGGCTGATCAACTGGGAGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1315
QY 2502 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2561
Db 1316 GAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1375
QY 2562 CTTGGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2621
Db 1376 CTTGGGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1435
QY 2622 TTCCCGTGCATCCGAAAGAGGCTGACCTCAAGGACACACCTGCTGCTGCTGCTGCTGCTGCT 2681
Db 1436 CTTCCGAGTGCATCCGAAAGAGCTGACCTCAAGGACACACCTGCTGCTGCTGCTGCTGCTGCT 1495
QY 2682 CACCTTCAATTTTGAACCGGCTGATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2741
Db 1496 CACCTTCAATTTTGAACCGGCTGATGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1555
QY 2742 CGCCAGGCTTGGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2801
Db 1556 CACCAAGGCTTGGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1615
QY 2802 TGAGTTCACCTTCCGAGGCTTCTGATGACAGGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCT 2861
Db 1616 CGAATTCACCTTCCGAGGCTTCTGATGACAGGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCT 1675
QY 2862 CTCAG-----AGCCCGCAGAGTGTGACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2909
Db 1676 TTCAGAGCAGAGGCTTTCAGAGCTGAGAGCTGAGATTTAGAAAGCTGCTGATCTGCTGCTGCTGCT 1735
QY 2910 TGGGCGGCGGCTGCTGAGGCTGAGAGCTGAGAGCTGAGATTTAGAAAGCTGCTGATCTGCTGCTGCT 2942
Db 1736 TCAGACTGAAGGCTGCTGAGGCTGAGAGCTGAGAGCTGAGATTTAGAAAGCTGCTGATCTGCTGCTGCT 1768

```

RESULT 3

```

US-09-509-902A-6
; Sequence 6, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-6

```

```

Query Match      36.4%; Score 1148.4; DB 4; Length 1498;
Best Local Similarity 86.3%; Pred. No. 8.6e-293;
Matches 1295; Conservative 0; Mismatches 196; Indels 9; Gaps 2;

```

```

QY 1185 CCCGCTGGATGTGGAGTATCGGGCTCCGAGCGCAGGTGCGCCACCGCTGGACCTGGAGAGCTGTCAC 1244
Db 4 CCCGCTGGAGCTGGAGTACGGAGGCGCCAGACCGGAGGTGCGCCCTCCGCTACCCGAA 63
QY 1245 GCATCTGCTGCTGCCCAAGTAACTGTGAGCAGTACACGCTGGACCTGGAGAGCTGTCAC 1304
Db 64 GCACCTGCTGCTGGCAGCAAGTCGGAGCAGTAC-----GACCTGGAGAGCTGTCGCG 117
QY 1305 CAGTGTGACGAGCTGCTGGAGGGG---CACTGATCTAGACGGGAGTGCACAAAGAGCA 1361
Db 118 AGCATGTGAGCAGAGCTGCTGCGGGCCCAAGAGCGCCGAGGGGCGGAGCAAGAGCG 177
QY 1362 CAAAGGTGCGAAGGAGACAAAGCTGGCAGAGACAAAGAGAGATTCAGACCTCCCGGT 1421
Db 178 CAAAGCGCCAGGGGACAAAGCGGAAAGGATAAAAGAGAGATTCAGACCTCCCGGT 237
QY 1422 GCTGTGCGCAAGATPAGCAGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTACTC 1481
Db 238 TCCGCTCCGCAAAACAGCAGAGACAGAGAGAGAGAGTCAAGCATCAAGAGTACTC 297
QY 1482 CCTTATGCTTCAATTTCTTCATGGAGCAACAGCTGGAGAGATGTCAATCAAACTACCA 1541
Db 298 GCATACGCTTTAAGTTCTTCATGGAGCAGCAGCTGGAGAGATGTCAATCAAACTACCA 357
QY 1542 GCAGAGCTGACCGCGAGGCTACAGCTGGAGCAGGAAATGCCAAAGCTGGGCTCTGTA 1601
Db 358 GCAGAGGTTTAACTGCGAGGCTGACGCTGGAGCAAGATGCCAAAGCTGGAGCTCTGTA 417
QY 1602 GCGCAGCAGGAGCAGATGAGAGATCCTTACACAGAGGAGTCTAACTACAACTGCT 1661
Db 418 AGCTGAGCAGGAGCAGATGCGGAGAGATCCTTACAGAGAGAGTCTAATTAACAGGTT 477
QY 1662 GAAGAGGCGCAAGATGGACAAGTCCATGTTGTGAATCAAGACTTAGGCATCGGTGC 1721
Db 478 AAAGAGGCGCAAGATGGACAAGTCTATGTTGTCAAGATCAAACTGCGGAGTCTGTCG 537
QY 1722 CTTTGGGAGTGTGCTGCTGCTTGAAGCTGGACACTCAGGCTCTGTACGCCATGAAGAC 1781
Db 538 CTTTGGGAGTGTGCTGCTTGTGAAGTGGACACTCAGGCTCTGTACGCCATGAAGAC 597
QY 1782 TCTCAGGAGAGAGATGCTGCTGAGCCGAGATCAAGTGGGCCCATGTCAAGCTCAGAGGGA 1841
Db 598 CTTAAGGAGAGAGATGCTGCTGAGCCGAGATCAGGTGGGCCCATGTCAAGCTCAGAGGGA 657
QY 1842 CATCTCGCTGAGAGAGAGATGAGTGGGTGGTCAAACTTACTACTCTCTCTCTCTCTCTCTCTCT 1901
Db 658 CATCTCGCTGAGAGAGAGATGAGTGGGTGGTCAAACTTACTACTCTCTCTCTCTCTCTCTCTCTCT 1901

```

QY 1902 GGACAGCCTGTACTTTGTGATGACTACATACAGCGGGGATATGATGAGCCTGTGTAT 1961
 Db 718 ACACAGCCTGTACTTTGTGATGACTACATCCTGTGGGACATGATGAGCCTGTGTAT 777
 QY 1962 CAGGATGAGGCTTCCCTGACACCTGGCCGGCTTACATTTGAGAGTTGACCTGGC 2021
 Db 778 CCGGATGAGGCTTCCCTGACACCTGGCCGGCTTACATTTGAGAGTTGACCTGGC 837
 QY 2022 CATTGAAAGTGTCCACAGATGGCTTTATCCACGGGACATCAAGCCTTGACCAATACT 2081
 Db 838 CATTGAGAGTGTCCACAGATGGCTTTATCCACGGGACATCAAGCCTTGACCAATTT 897
 QY 2082 CATGACCTGGATGTATTAAGCTGACAGATTTTGGCCCTCTGACCTGGATTCAGGTG 2141
 Db 898 GATGATCTGGATGTCTCACTAACTCAACATTTGGCCCTCTGACCTGGGTTCAGGTG 957
 QY 2142 GACTCACAATTCAGTACTACCAAGAGGGAACACATGACAGGACGATGAGCC 2201
 Db 958 GACTCACAATTCAGTACTACCAAGAGGGAACATGACAGGACGATGAGCC 1017
 QY 2202 CGGTGACCTCTGGGACGATTTTCAACTGTGCTGTGGAGACAGTTAAAGACCTTGA 2261
 Db 1018 CAGCAGCTCTGGATGTCTTAAGTGTGGGTGGGACAGCTGAAGACCTTGA 1077
 QY 2262 GCAGAGGGCGGAGAGCAGCAGAGGTGCTTGGACATTTCTTTGGGACACCAAA 2321
 Db 1078 GCAGAGGGCGGAGAGCAGCAGAGGTGCTTGGACATTTCTTTGGGACACCAAA 1137
 QY 2322 TTACATCGCTCGGAGGTGCTTCCGCAAGAGGTACAGGAGCTGTGACTGTGGAG 2381
 Db 1138 CTACATCGCACCAGGTGCTTCCGCAAGAGGTACAGGAGCTGTGACTGTGGAG 1197
 QY 2382 COTCGGTGTATTTCTTTGAGATGCTGTTGGGACGCGCTTTCTTTGGCCCGCACCC 2441
 Db 1198 TGTGGAGTATCTTCTCGAGATGCTGTTGGGACGCGCTTTCTTTGGCCCGCACCC 1257
 QY 2442 CACAGAGCGCAGCTGAAGTGTATCACTGGGAGAGCAGCTGTGATATCCCTTACGAGT 2501
 Db 1258 CACAGAAACCCAGCTGAAGTGTATCACTGGGAGAGCAGCTGTGATATCCAGCCAGT 1317
 QY 2502 GAGCTCAGCTGAGGCGGAGAGCTTATCAGAGTGTGCTGCGGCGGTGACTGGC 2561
 Db 1318 GAAGTGAAGCCTGAGGCGGAGAGCTTATCAGAGTGTGCTGCGGCGGTGACTGGC 1377
 QY 2562 COTGGCAGGAGTGGGAGAGTGTATCAAGCAGCAGCTTCTTCAACACAGCTGACTT 2621
 Db 1378 COTGGCAGGAGTGGGAGAGTGTATCAAGCAGCAGCTTCTTCAACACAGCTGACTT 1437
 QY 2622 TTCCCGTGACATCCGAAAGCAGGTGACCTTACCTTACCTTACCTTACCTTACCTT 2681
 Db 1438 CTCCAGTGACATCCGAAAGCAGGTGACCTTACCTTACCTTACCTTACCTTACCTT 1497

RESULT 4
 US-09-442-100-5
 ; Sequence 5, Application US/09442100
 ; Patent No. 6359193

GENERAL INFORMATION:

APPLICANT: Xu, Tian
 APPLICANT: Tao, Wufan
 APPLICANT: Wang, Weiyl
 APPLICANT: Zhang, Sheng
 APPLICANT: Yu, Wan

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/442.100
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/411.111
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Mirock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6523-003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3213 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..2889
 US-09-442-100-5

Query Match 25.8%; Score 812.6; DB 4; Length 3213;
 Best Local Similarity 65.1%; Pred. No. 3.4e-204;
 Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTACTGCCGCCCAACACGCGTCAACGCGTCAACGCGCCGACACATC 1044
 Db 919 CACTCTGTCTTAATCTCAGCCCTTCTGCCACTCAGTCCATCACACCCGCTCTATT 978
 QY 1045 CTTACCTCTGTAAGAGCGTGGTGTGTCGGCCGCGCCGACAGCCGCTGGGGGCC 1104
 Db 979 CAACAGCCGCTGAAAGCATGCGCTCTCTGAAACACAGAGCTCGACTGCTTTAGCCCA 1038
 QY 1105 TCGCACCCGCGCTGCGCCGCGTCCGCTGCGTCCGCTGCGTCCGCTGCGTCCGCTGCG 1164
 Db 1039 ACCATCTCTTGTGATGCCACAGCCAGTTTACAGCTGTTCAGCTTACCCCTTTTCTGAG 1098
 QY 1165 GAGGCGAGCGAGCGCCACACCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGC 1224
 Db 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTGTGCTGAAGCTCCAAGCTATCAAGT 1158
 QY 1225 CCACCGCTCCGTATCCAAAGCAGTTCGTGTCGCCAGTAACTGAGCTGAGCAGTACAGCTG 1284
 Db 1159 CCACCGCTCCGTATCCAAAGCAGTTCGTGTCGCCAGTAACTGAGCTGAGCAGTACAGCTG 1218
 QY 1285 GACCTGGAGCAGCTGTGCGACAGTGTGCGAGAGAGTTCGAGGGGCGCAGTGTAGTACAG 1344
 Db 1219 TCAGTAACTAGCCCTGCAAGATGAACAGCTTACCCCAAGAGATGATAGTACAG 1278
 QY 1345 GGGAGTGACAGAGCCACAAAGGTGCGAAGGAGAGACAAAGCTGGCAGAGACAAAAGCAG 1404
 Db 1279 AAGAGTGGGACAGT-----GGTGACTCTGGGATATAAGAAAGAAACAG 1323
 QY 1405 ATTCAGACTTCCCGGTGCTGTCGCAAGATATACAGAGATGAAGAGAGAGAGTCT 1464
 Db 1324 ATTACAACTTCACTTATCACTGTTTCGAAACAAAGAAAGATGAAGAGAGAGTCT 1383
 QY 1465 CGCATCAAGAGTACTCCCTTATGCTTCAATTTTCAATTTTCAATTTTCAATTTTCAATTTT 1524
 Db 1384 CGGATTCAGAGTTACTCCCGACAGGCTTTTAAAGTTTCTTCTTCTTCTTCTTCTTCTTCT 1443
 QY 1525 GTCATCAAAACCTTACCAGCAAGAGGTTCAGCGGAGGCTTACAGCTGGAGCAGGAAATGGCC 1584

2524	DB	TTTAAGACCATCGATTTTCTCTAGTGATCTGAGACGAGCTGCTGCTTCATACATCCCTAAA	2583
2665	QY	ATCAGCGACCCCATGGACACCTCCAAATTTTCACCCGGTGGNTGAGAAAGCCCTCGCAC	2724
2584	DB	ATCAGCGATCCAACAGATACATCCNATTTTCAGCCCTGTTGATCTGTATAAATTGTGGAGC	2643
2725	QY	GAGGCCAGCGGAGAG --- AGCGCCCAAGCGCTGGACACGCTGGCGCTCCCCAGCAGCAAG	2781
2644	DB	GATGCGACGGAGGAGGAAAAATATCAGTCACACTCTGAGCGGATGGTATAAAAAATGGGAAG	2703
2782	QY	CATCCAGACGACGCGCTTCTATGAGTTCACCTTCGCGAGGTTCTTCGATGACAGCGCCTAT	2841
2704	DB	CACCCCGACGACGCTTCTATGAGTTCACCTTCGAGGTTTTTGATGACAATGGCTAC	2763
2842	QY	CCCTTCGGTGCCCGAAGCCCTCAGAGCCCGCAGAGAGT	2880
2764	DB	CCATATAATATCCAAAGCCTATTGAGTATGAATACATT	2802

RESULT 5
US-09-442-100-3
; Sequence 3, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LAT5
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON

Query Match	24.1%;	Score 760;	DB 4;	Length 3984;
Best Local Similarity	63.7%;	Pred. No. 2.6e-190;		

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTCACTGCCGCCGCCCAACAGCGGTACCGCGTACGCCCGCACACATC 1044
 Db 1653 CACTCTGCTAAATCTCAGCGCTTCTGCTACACAGTCACTGCAATTACACCGACTCCTATT 1712
 QY 1045 CTTTCACTCTGTAAGAGCGTGTGCTGTGCGCGCGCGAGCCCGACAGACCGCTGGGGCC 1104
 Db 1713 CAACAGCGCTGTGAAAGTATGCGTGTATTAACACAGAGCTACAGACTGCTTTAGCACCT 1772
 QY 1105 TCGCACCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1164
 Db 1773 ACACACCTCTCTGGTATACAGACCAATTAACAGTGTTCACACCGACTCTTTTCTGTAG 1832
 QY 1165 GAGGAGCAGCGCAGCGCCACACCGCTGGTATGTGGACTATGGCGGTCTCCGAGCGCAGGTGC 1224
 Db 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTTCCAACTATCAAGGA 1892
 QY 1225 CCACCGCTCTGCTATTCAAAGCACTTGTGCTGCGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
 Db 1893 CCACCAACCGCTTACCCCAAAACATCTGTGCAACCAACCCCATCTGTTCCTCCATACGAG 1952
 QY 1285 GACCTGGACAGCTGTGCACCACTGTGCAGCAGAGTCTGCAGGGGGCAGCTGATCTAGAC 1344
 Db 1953 TCAATCAGTAAGCTAGCAAGAGGATCAGCCAGCTTGCACCAAGAGATGAGAGTAA 2012
 QY 1345 GGGAGTGAAGAGCCACAAAGGTGCGAAGGAGAGCAAGCTGGCAGAGACAAAAGCAG 1404
 Db 2013 AAGAGTTA-----TGAATGTTGATAGTGGGGATTAAGAAAGAAACAG 2057
 QY 1405 ATTCAAGCTCCCGGTGCTGCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 Db 2058 ATTCAAGCTTCACTTCTTCTCAAGCACTTAAATTTCTTTTGGAGCAACATGAGAAAT 2117
 QY 1465 CGCATCAAGAGTTACTCCCTTATCCCTTCAATTTCTTCATGCGACCAACAGTGGAGAT 1524
 Db 2118 CTATTCAAGTTATCTCTTCAAGCACTTAAATTTCTTTTGGAGCAACATGAGAAAT 2177
 QY 1525 GTCAATCAAAACCTACAGCAGAGGTGACCGCGAGGTACAGCTGGAGCAGGAAATGGCC 1584
 Db 2178 GTACTCAATCTCATCAGCAGCGTCTACATCGTAAACCAATTAGAATGATGATG 2237
 QY 1595 AAAGCTGGCTGTGTAGCGCGAGCAGGAGCAGATGAGGAGATCCTTACCAAGAGG 1644
 Db 2238 CGGTTGAGTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 2297
 QY 1645 TCTAACTACAACCGCTGAAGAGGCGCAAGATGGACAGTCCATCTTCTGGAATCAAG 1704
 Db 2298 TCTAAATACATCCGCTTAAAGGGCTAAATGGACACATCTATGTTTGTGAAGATAAG 2357
 QY 1705 ACTTAGCATCGGTGCTTTGGGGAAGTGTGCTCGCTTGTAAAGCTGGACACTCACGCT 1764
 Db 2358 ACATPAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATAAGGCT 2417
 QY 1765 CTGTAGCCATGAAGACTCTCAGGAAGAGATGTCCTGACCGGAATCAAGTGGCCCAT 1824
 Db 2418 TTGTATGAACAAAACCTCTTCGAAAGAAAGATGTTCTTCTTCGAAATCAAGTGCCTCAT 2477
 QY 1825 GTCAAGGCTGAGGGAGACATCTGTGCTGAAGCAGACATGAGTGGTGGTCAAACTCTAC 1884
 Db 2478 GTTAGGCTGAGAGAGATATCTGTGCTGAAGCTGAACATGAATGGTGTGCTGTATAT 2537
 QY 1885 TACTCTTCCAGGCAAGCAGCAGCTGTACTTGTGTGAGTACTACATACAGCGGGGAT 1944
 Db 2538 TATTCAATCCAAGATAAGCAATTTATCTTGTATGAGTACTATCTTCTGGGGGTGAT 2597
 QY 1945 ATGATGAGCTGTGATCAGGATGAGTCTTCCCTGAGCAGCTTCCCGCGCTTCTACATT 2004
 Db 2598 ATGATGAGCTTATTAATTAAGATGGCATCTTCCAGAAAGTCTGGCAGCATTTCTACATA 2657
 QY 2005 GCAGAGTTGACCTGGCCATGAAGTGTCCACAGATGGCTTTATCCACCGGGACATC 2064
 Db 2658 GCAGAACTTACCTGTGCAGTTGAAGTGTTCATAAAATGGTGTATTTATCATAGAGATAT 2717

QY 2065 AAGCCTGACAACATACTCATCGACCTGGATGCTATATTAAGCTGACAGATTTGGCCCTC 2124
 Db 2718 AAACCTGATAATATTTGATTGATCGTATGATGCTATATTAATTAAGTACTGACTTTGGCCCTC 2777
 QY 2125 TCACTGGATTCAGGTGGACTCACAATTCAAAGTACTACAGAAAGGGGAACACATAGAGA 2184
 Db 2778 TCACTGGCTTCAGATGGACACACAGATCTAAGTACTATCAGAGTGTGACCATCCACGG 2837
 QY 2185 CAGGACAGCATGGACCGCGCTGACCTCTGGGAGGATGTTTCCAACTGTGCTGTGGAGAC 2244
 Db 2838 CAAGATAGCATGGATTTCAATGAGGGGGATCCCTCAAGCTGTGCTGTGGAGAC 2897
 QY 2245 AGGTTAAACACCTCGAGCAGAGGGCGCAAGAGCAGCAGCAGAGGTGCTGCACATTTCT 2304
 Db 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGCACCCAGCAGCAGGATGTCTACACATTTCT 2957
 QY 2305 CTTTGGGACACCAAAATTTACATCCTCGGAGGTGCTTCTCCGAAAGGGTACACGAG 2364
 Db 2958 TTGGTTGGGACTCCCAATTTATTTGACCTGAAGTGTGCTACGAACAGGATACACAG 3017
 QY 2365 CTCCTGAGTGTGAGCGTGTGATGTTCTTTGAGATGCTGTGGGAGCGGCT 2424
 Db 3018 TTGTGTGATTTGGTGGAGTGTGTTGTTTGTGAAATGTTGGTGGGACAACTCTCT 3077
 QY 2425 TTCTTTGGCCCCCACCACACAGAGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTG 2484
 Db 3078 TTTCTTGGCACAACACCATTTAGAAACACAAATGAAGTTATCACTGGCAACATCTCTT 3137
 QY 2485 CATATCCCTTACCGAGTGTGAGCTCAGCGCTGAGCGCCGAGACCTCATCAGAAAGTGTGC 2544
 Db 3138 CACATTTCCACCAAGCTAAACTCAGTCTGAAAGCTTCTGATCTTATTATTAACTTTGC 3197
 QY 2545 TCGCGGCTGACTCGCGCTGGGAGGATGGGCGAGATGAGCTCAAGGCACACCGCTC 2604
 Db 3198 CGAGGACCCGAAAGTCGCTTAGGCAAGATGTTGATGAAATAAAGCTCATCCATTT 3257
 QY 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGCAGCTGCACCTTACGTCGCCACC 2664
 Db 3258 TTTAAACAATGTACTTCTCCAGTCACCTGACGACAGCAGTCTGCTTACATTTCTTAA 3317
 QY 2665 ATCAGCCACCCCATCGACACTTCAATTTTACCGGTTGATGAAGAAAGCCCTGGCAC 2724
 Db 3318 ATCAGCACCACCAAGATACATCAATTTTATCTGTTGATCTCTGATAAATATTAGGCT 3377
 QY 2725 GAGGCCAG- --GGAGAGCGCCAGGCTGGGACAGCTGGCTTCCCGCAGCAGCAG 2781
 Db 3378 GATGATAAGAGGAAAGAAATGAAATGACACTCTCAATGGATGATATAAAATGGAAG 3437
 QY 2782 CATCCAGAGCAGCGCTTCTATGATGTTCCCTTCCGAGCTTCTTCGATGACACCGCTAT 2841
 Db 3438 CATCTGAACATGCTTCTATGAAATTTACCTTCCGAGGTTTTTGTGATGACATGGCTAC 3497
 QY 2842 CCCTTCCGGTGGCCGAGGCC 2861
 Db 3498 CCATATAATTATCCGAGGCC 3517

RESULT 6

US-09-442-100-1
 ; Sequence 1, Application US/09442100
 ; Patent No. 6359193
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Tian
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Wang, Weiyl
 ; APPLICANT: Zhang, Sheng
 ; APPLICANT: Yu, Wan
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 ; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds

RESULT 7

US-09-328-111-26/c
; Sequence 26, Application US/09328111
; Patent No. 6262333

GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert

TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

TITLE OF INVENTION: PRODUCTS

FILE REFERENCE: CCD-257 (US)

CURRENT APPLICATION NUMBER: US/09/328,111

CURRENT FILING DATE: 1999-06-08

EARLIER APPLICATION NUMBER: US 60/088,801

EARLIER FILING DATE: 1998-06-10

NUMBER OF SEQ ID NOS: 850

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 26

LENGTH: 638

TYPE: DNA

ORGANISM: Homo sapiens

US-09-328-111-26

Query Match 10.6%; Score 333; DB 4; Length 638;

Best Local Similarity 71.3%; Pred. No. 3.4e-78;

Matches 454; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1525 GTCATCAAAACCTACCAGCAGAGGTGACGGGAGGCTACAGCTGGAGCAGGAATGGCC 1584

DB 637 GTACTCAAACTCTACAGCAGCGTCTACATCGTAAAAACAATAGAGAATGAATGATG 578

QY 1585 AAAGCTGGGCTCTGTGAGCGGAGGAGGAGGAGATGAGGAGATCCTCTACGAGAGAG 1644

DB 577 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAA 518

QY 1645 TCTAATCAACCGGCTGAAGAGGCCAAGATGGAAGTCCATTTGTGAAATCAAG 1704

DB 517 TCTAATACATCCGCTTAAAGGGCTAAATGAGCAAGTCTATGTTGTGAGAGTAAG 458

QY 1705 ACTTAGCATCGGTGCGCTTTGGGAGTGTGCGCTGCTGTAAGCTGGACACTCACGCT 1764

DB 457 ACACATAGATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACCTAAGCT 398

QY 1765 CTGTACCCATGAAGACTCTCAGAGAGAGATGCTCTGACCGGAATCAAGTGGCCCAT 1824

DB 397 TTGTATGCAACAAAACCTCTCGAAATAAAGATGTTCTTTCGAAATCAAGTCCGCTCAT 338

QY 1825 GTCAAGGCTGAGGGAGCATCTGTGCTGAAGCAGACATGAGTGGTGGTCAAACTCTAC 1884

DB 337 GTTAAGCTGAGAGATATCTGTGCTGAAGCTGACATGAATGAGTGTGCTGTATAT 278

QY 1885 TACTCTCTCCAGGCAAGGAGCAGCTGTACTTTGTGATGAGTACATACACGAGGGGAT 1944

DB 277 TATTCTTCCAAAGTAAGCAATTTATATCTTTGTAATGGACTACATTCCTGGGGTGTAT 218

QY 1945 ATGATAGCTGTGATCAGATGAGGTCTTCCCTGAGCAGCTGGCCCGCTTCTACATTT 2004

DB 217 ATGATAGCTTATTAATTTAGAAATGGCATCTTCCAGAAAGCTGGCAGCATTTACATA 158

QY 2005 GCAGATTTGACCTGGCCATTTGAAGTG---TCCACAGATGGGCTTTTATCCACGGGAC 2061

DB 157 GCAGACTTACTGTGAGTTTGAAGTGTCTTTAAATGGGTTTTATTTCATAGAGAT 98

QY 2062 ATCAAGCCTGACAACTACTCATCGACCTGGATGGTGTATATTAAGCTGACAGATTTTGGC 2121

DB 97 ATTAACCTGATAATATTTTGTATCGTGTATGCTATATAATTGACTGACTTTGGC 38

QY 2122 CTCTGCACTGGATTCAGGTGGACTCACAATTCCAAGT 2158

DB 37 CTCTGCACTGGCTTCAGATGGACACACGATTTCTAAGT 1

RESULT 8

US-09-588-256-1

; Sequence 1, Application US/09588256

; Patent No. 6291665

GENERAL INFORMATION:

; APPLICANT: Gaffney, Thomas

; APPLICANT: Flaviey, Albert

; APPLICANT: Gates, Krista

; APPLICANT: Wendland, Juergen

; APPLICANT: Ayad-Burieux, Yasmina

; APPLICANT: Dietrich, Fred

; APPLICANT: Philippsen, Peter

TITLE OF INVENTION: Fungal Target Genes and Methods

FILE REFERENCE: PB/5-30908A

CURRENT APPLICATION NUMBER: US/09/588,256

CURRENT FILING DATE: 2000-06-06

NUMBER OF SEQ ID NOS: 24

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1

LENGTH: 2160

TYPE: DNA

ORGANISM: Ashbya gossypii

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2160)

US-09-588-256-1

Query Match 8.2%; Score 258.8; DB 4; Length 2160;

Best Local Similarity 54.6%; Pred. No. 2e-58;

Matches 598; Conservative 0; Mismatches 447; Indels 51; Gaps 2;

QY 1637 AGAAGGAGCTTAATCAACCGGCTGAAGAGGGCCCAAGTGGACAGTCCATGTTTGA 1696

DB 875 AAAAGAGTCGCGATTTCTGCGCTTTGCGTAGGACACGCTATCCCTGGGAAGATTTCCACA 934

QY 1697 AAATCAAGACTCTAGCATCGTGTGCTTTGGGAGTGTGCTGCTGTAAGCTGGACA 1756

DB 935 CTGTAAAGTATAGGAAGGGTGCATTCGCTGAGTCCGCTGCTGTCAGAGAAGATA 994

QY 1757 CTCAGCTCTGTAGCCCATGAAGACTCTCAGGAAGAAGATGTCTCTGAACCGGAATCAAG 1816

DB 995 CCGGTAATAATATACGCTATGAAGACATTTTAAATCAGAAATGTACAAGAAGATCAAT 1054

QY 1817 TGGCCCATGTCAAGGCTCAGAGGGACATCTCGGCTGAAGCAGACATGAGTGGTGGTCA 1876

DB 1055 TAGCCAGCTCAAGCGCGAGAGGGATGTGTGGCGGGAAGCAGCTCTCCGTGGGTGCTGT 1114

QY 1877 AACTCTACTCTCTCCAGGCAAGGACAGCTCTACTTTGTGATGGACTACATACACAG 1936

DB 1115 GGTATATATCTTCCAGATGCCAGTACCTATATCTGATCAATGGAATTTTGGCCG 1174

QY 1937 CGGGGATATGATAGCTGTGATCAGGATGGAGGTCTTCCCTGAGCAGCTGGCCCGCT 1996

DB 1175 GTGGTGACCTGATGACCATTTAATCAGGTGGCAGATTTTCCAGGAGCAGCTCACCAT 1234

QY 1997 TCTACATTTGAGATTTGACCTGGCCATTTGAAGTGTCCACAGATGGCTTTATCCACC 2056

DB 1235 TCTACATCGCGAGTGTATCTGCAATTTAGGCTTATACAAAGCTGGGCTTTATCCATA 1294

QY 2057 GGGACATCAAGCTGACAACTACTCATCGACCTGGATGGTGCATATTAACTGACAGATT 2116

DB 1295 GAGATATCAAGCCGATTAACATTTCTGATCAGATCAGGGGTCAATCAAACTTTCCGACT 1354

QY 2117 TTGGCTCTGCACTGGATTTCAAGTGGACTCAATTTCCAAGTACTTACCAGAAAGGGAACC 2176

Fri Jan 17 11:17:04 2003

```

1355 TCAGTCTGTGACAGGTTCCCAAAACGATGACTTCCAACTACTACAAGAGCTGCTTC 1414
QY 2177 ACATGAGACAGGACAGCATGGAGCCCGGTGACCTCTGGAGCATGTTTCCAACTGCTGCT 2236
Db 1415 AGGAGGACAGGACAGCATGGAGCCCGGTGACCTCTGGAGCATGTTTCCAACTGCTGCTG 1474
QY 2237 GTGGAGA-----CAGGT 2248
Db 1475 GCGGCAACGCGGCGGCAACAGCAACACCATGCTTCTGACGCGCATCCACCTGACCATGA 1534
QY 2249 TAAAGACCTTGGAGAGGCGGCAAGCAGCAGCAGAGGTCCTTGGGACATCTCTTGG 2308
Db 1535 CARACAGGACAGCATGCAAACTGCGCAAGTCCGCTAGGCTCATGGGCTACTTCCACCG 1594
QY 2309 TCGGAGACCAAAATTACATCGCTCCGAGGCTTCTTCCGCAAGGTTACAGCAGCTCT 2368
Db 1595 TCGGTACGCCAGACTACATCGCCCGGAGATCTTCTTACAGGAGGCTACGGTTCAGGAGT 1654
QY 2369 GTGACTGTGTGAGGCGGTGTGATTCTCTTTGAGATGCTGTGGTGGCAGCCGCTTCT 2428
Db 1655 CGACTGTGTGCTTCCGCGCCCATCATGTACGAGTGTCTGTGATCGGCTGGCCGCTTCT 1714
QY 2429 TGGCCCCCACCACACAGAGCAGCAGCTGAGGTGATCACTGGGAGCAGCAGCTGCATA 2488
Db 1715 GCTGGAGACCCCGGCAAGACCTTACAGGAAGATCATGAATCTGAGCAGCAGCAGCTGTGT 1774
QY 2489 TCCCTTACGAGTGTGAGGCTCAGCGCTGAGGCGGAGACCTCATCATCAAGAGCTGTGCTGG 2548
Db 1775 TCCAGAGCAGCATCCACATCTCATAGGAGCAGGAGGACCTCATCCGCGGCTGTCTCGC 1834
QY 2549 CGGCTGACTGCGGCTGGGAGGAGTGGGAGGATGACCTCAAGGCAACACCGCTTCTTCA 2608
Db 1835 AGCGGAGAGCAGGCTGCGGTGCGGATGCGGCAAGCAAGCAATCAAAACCAACCCCTTCTTC 1894
QY 2609 ACACATGACATTTTCCGCTGACATCCGAAAGCAGGCTGACCCCTAGCTGCCACCATCA 2668
Db 1895 GCGGCTGGAC---TGGGAGACCATCCGCGGCTCCCTACATATCCCAAGCTGT 1951
QY 2669 GCCACCCCATGAGAC 2684
Db 1952 CCAGCGTCCAGGACAC 1967

```

RESULT 9

US-08-878-989-11
Sequence 11, Application US/08878989
Patent No. 5885803

GENERAL INFORMATION:

APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,989
FILING DATE:

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1935 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT04
CLONE: 705365
US-08-878-989-11

```

Query Match 7.4%; Score 233.8; DB 2; Length 1935;

Best Local Similarity 51.7%; Pred. No. 7.4e-52;

Matches 609; Conservative 0; Mismatches 532; Indels 36; Gaps 2;

```

QY 1494 CAAATTTCTTCATGGAGCAACACGCTGGAGAAATGTATCAAAACCTTACCAGCAGAGGTCAG 1553
Db 121 CAAGCTCACATTTGGAGAAATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAGAAAC 180
QY 1554 CGGAGGCTACAGCTGGAGCAAGAAATGCCAAAGCTGGGCTCTGTAGGCGCCGAGCAGGA 1613
Db 181 CAGGCAAGAGAAATAGAAAGTGGCCATGGAAGAGAGAGATTAGCAGATGAAGAGAAAA 240
QY 1614 GCAGATGAGGAAGATCCTCTTACCAGAGAGAGTCTTAACATAACCGGCTTGAAGAGGCCAA 1673
Db 241 GTTACGTCGATCACACACGCTCGCAAGAAACAGAGTTCTTACGGCTCAAAAGGACGAC 300
QY 1674 GATGGACAGCTCCATGTTTGTGAANAATCAAGACTCTAGGCACTCGGTGCTTTGGGGAAGT 1733
Db 301 ACTTGGCTTGGATGACTTTTGGATCTCTGAAAGTTATAGAGAGGAGCTTTTGGAGAGT 360
QY 1734 GTGCCTCGCTTGTAAAGCTGGACACTCACGCTCTGTAGCCATGAAGACTCTCAGGNAGAA 1793
Db 361 CCGGTTGCTCCACAAAAAGATACAGGCCATATCTATGCAATGAAGATATTGAGAAAGTC 420
QY 1794 GGATGCTCTGAACCGGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCTGGCTGA 1853
Db 421 TGATATGCTTGAAGAGAGCAGGTGGCCCATATCCGAGCAGAGAGAGATATTTTGTAGA 480
QY 1854 AGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGCAGAGAGAGAGAGCTGTA 1913
Db 481 AGCAGATGGTGGCTGGTGGTGGTGGATGTTTACAGTTTTCAGCATGAAGAGAGATCTTTA 540
QY 1914 CTTTGTGATGAGCTACATACAGCGGGGATATGATGAGCCTGCTGATCAGGATGAGAGT 1973
Db 541 TCTAATCATGGAATTTCTCCCTGGAGGTGATGATGATGATGATGATGATGATGATGATGATG 600
QY 1974 CTTCCCTGAGCACCTTGGCCGCTTCTACATTCAGAGTTCACCTTGCCCTGAGAGTGT 2033
Db 601 CTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 2034 CCACAAGATGGGCTTTATCCACCGGAGATCAAGCTTCAGACATCTATCTATCTATCTATCT 2093
Db 661 CCACCAAGTGGGTTTTCATCCATCGGGATATTAAGCAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 2094 TGGTCATATTAGCTGACAGATTTTGGGCTCTGCACTGGATTCAGGTCAGGTCACAATTC 2153
Db 721 GGGTCATGTAATAATTTATCTGATTTTGGTTCATGTACGGGATTTAAAGAGAGAGAGAG 780
QY 2154 CAAGTACTACCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2213

```

Db 781 TGAATTTATAGAAATCTCACACACAAACCCACCA-----AGTGACTTCTCATTTTC 830
QY 2214 GGACGATGTTTCCAACTGTCGTGGAGACAGGTTAAAGACCTTGGAGCAGAGGCGCA 2273
Db 831 AGAACATGAATCTCAAGAGAAAGCAGAACTTG----- 864
QY 2274 GAAGCAGCAGAGGTCCTGGCAGATCTCTTGTGGGACACCAAAATTACATCGCTCC 2333
Db 865 GAAGAAGAACAGGACAGCAACTGGCATATTCACAGTTGGGACACCAAGATTACATTCCTCC 924
QY 2334 GGAGTGCTTCTCCGCAAGAGGTACACGAGCTCTGTGACTGTTGGAGGCTGGTGTGAT 2393
Db 925 AGAAGTATTCAGCAGACTGGTTACAAATTTGTGACTGTTGGGAGTGAT 984
QY 2394 TCTCTTTGAGATGCTGGTGGCAGCGCTTCTTGGCCGCCACCCACAGAGACGCA 2453
Db 985 TATGTATGAATGCTAATAGGATATCCACTTTCGCTGAAACACCTCAGAAACATA 1044
QY 2454 GCTGAAGGTGATCACTGGGAGACGCTGCATATCCCTACGCGAGGTGAGCTCAGCGC 2513
Db 1045 CAGAAAAGTGATGAATGGAAGAAATCTGTGATTTCTCCAGAGGTACCTATATCTGA 1104
QY 2514 TGAGGCCGAGACCTCATCACAAAGCTGTGCTGCGGGCTGACTCGCGCTGGGAGGA 2573
Db 1105 GAAAGCCAGGACTTAATCTCAGATTTTGTATTGATTCTGAAACACAGAAATGGAATAG 1164
QY 2574 TGGGCGATGACCTCAAGGCACACCGTTCTTCAACACCATCGACTTTTCCCGTGACAT 2633
Db 1165 TGGAGTAGAGAAATAAAGGTTCATCCCTTTTGAAGGTGTCGACTGGGACCATATAG 1224
QY 2634 CCGAAGCAGGCTGCACCTAGCTCCGCCACCATCAGC 2670
Db 1225 GGAAGGCCAGCAGCAATCCCTATAGAAATCAAAAGC 1261

RESULT 10

US-09-272-796-11
; Sequence 11, Application US/09272796
; Patent No. 6207148

GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0321 US

TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAT04
; CLONE: 705365
; US-09-272-796-11

Query Match 7.4%; Score 233.8; DB 4; Length 1935;
Best Local Similarity 51.7%; Pred. No. 7.4e-52;
Matches 609; Conservative 0; Mismatches 532; Indels. 36; Gaps 2;

QY 1494 CAAATTTCTCATGGAGCAACACGCTGGAGAATGTCTCAAAACCTACCAGCAGAAGTCTAG 1553
Db 121 CAAGCTCAGATTTGGAGAATTTTATAGCACTAATTTTACAGCATGAAGAGAGAGAAAC 180
QY 1554 CCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTTGGGCTCTGTGAGGCCGACGAGCA 1613
Db 181 CAGGCAGAGAAATTAGAAGTGGCCATGGAAGAAGAGATTAGCAGATGAAGAGAGAAAA 240
QY 1614 GCAGATGAGGAAGATCTCTACCAGAAAGAGTCTTAACATAACCGGCTGAAGAGGGCCAA 1673
Db 241 GTTACGTCGATCACAAACACGCTCGCAAAAGAAACAGAGTTCTTACGGCTCAAAAGGCCAG 300
QY 1674 GATGACAAGTCCATGTTGTGAAAATCAAGACTCTAGCATCGCTGCTTTGGGGAAGT 1733
Db 301 ACTTGGCTTGGATGACTTTGAGTCTCTGAAAGTTATAGAGAGAGAGCTTTTGGAGAGGT 360
QY 1734 GTGCTCGCTTGTAACTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAA 1793
Db 361 GCGGTTGGTCCACAAAAAGATACAGGCCATATCTATCAATGAAGATATTGAGAAAGT 420
QY 1794 GGATGCTCTGACCCGATCAAGTGGCCCATCTCAAGCTGAGAGGGACATCCTGGCTGA 1853
Db 421 TGATATGCTTGAAGAGAGCAGGTGGCCCATATCCGACAGAAAGAGATATTTGGTAGA 480
QY 1854 AGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGACAAAGCAGAGCTCTGA 1913
Db 481 AGCAGATGCTGCTGGTGGTGAAGATGTTTACAGTTTTCAGGATAGAGAGAAATCTTTA 540
QY 1914 CTTTGTGATGACTATACATACCAGCGGGGATATGATGAGCCTGTGATCAGGATGGAGT 1973
Db 541 TCTAATCATGGAATTTCTCCCTGGAGGTGACATGATGACATTCGTTAATGAAGAAAGAC 600
QY 1974 CTTCCCTGAGCACCTGGCCCTTCTACATTCAGAGTTGACCTGGCCATTTGAAGTGT 2033
Db 601 CTTGACAGAGAGGAAACACAGTTCTACATTTTCAGAGACTGTTCTGGCAATAGATGCGAT 660
QY 2034 CCACAAAGATGGCTTTATCCACCCGGACATCAAGCTGACAACTACTCATCGACTCGA 2093
Db 661 CCACAGTTGGTTTCATCCATCGGATATTAAAGCCACACACCTTTTATGGATGCCAA 720
QY 2094 TGGTCATATTAAGCTGACAGATTTTGGCTCTGCACTCTGAGTTAGGTGAGCTACAATTC 2153
Db 721 GGGTCATGTAATAATTTATCTGATTTTGGTTCATGTACGGGATTAAGAGAAAGCTCACAGGAC 780
QY 2154 CAAGTACTACAGAAAGGAAACCATGACAGCAGCAGCATGGAGCCCGGTGACCTCTG 2213
Db 781 TGAATTTTATAGAAATCTCACACACACACCCACCA-----AGTGACTTCTCATTTTC 830
QY 2214 GGACGATGTTTCCAACTGTCGTGGAGACAGGTTAAAGACCTTGGAGCAGAGGCGCA 2273
Db 831 AGACATGAATCAAGAGAGAAAGCAGAACTTG----- 864
QY 2274 GAAGCAGCAGAGGCTGCTGGCAGATTTCTTGTGCGGACACCAAAATTACATCGCTCC 2333

Query Match	7.1%	Score 225.2;	DB.2:	Length 3018;
Best Local Similarity	52.2%;	Pred. No. 1.7e-493;		
Matches 577;	Conservative	0;	Mismatches 493;	Indels 36; Gaps 2;
1565	ACGTGGAGCAGGAATCGCCAAAGCTGGGCTCTGTGAGGCCGAGGACGAGACATGAGA	1624		
735	AGTTAGAAAGGTTCTGGACAGAGAGGCCCTAAAGATGAGGAGAAACGACTCCCGGAGAT	794		
1625	AGATCTCTACCGAGAGGAGTCTTAATCTACAACGGCTGAAGAGGCCACAGATGGACAAGT	1684		
795	CAGCACATGCTCGGAAGGAACACAGAGTTCTTCGTTTGAAGAGAACAGACTTGCATTGG	854		
1685	CCATGTTTGTGAAATCAAGACTCTAGGCATCGTGCCCTTTCGGGAAGTGTGCCTCGCTT	1744		
855	AAGATTTTGTAGTCCCTTAAAGTAATAGGCAGAGGACATTTGGTGAGGTACGGCTTTGTC	914		
1745	GTAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGTCCTGA	1804		
915	AGAAGAAGATACGGGACATGTGTATGCAATGAAATATCTTCGTTAAAGACGAGATGCTTG	974		

QY	1805	ACCGGAATCAAGTGGCCCATGCTAAGCGCTGAGAGGACATCTCGCTGAAGCAGACAAATG	1864
Db	975	AAAAAGAGCAGCGTTGGCCACATTCGTGCGGCGTGACATTCCTAGTGGAGCGACACGATT	1034
QY	1865	AGTGGGTGGTCAAACTCTACTACTCTCTCCAGGACAAGGACAGCGTACTTTGTGATGG	1924
Db	1035	TGTGGGTGTGAAAATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTACCTAATCATGG	1094
QY	1925	ACTACATACCAAGCGGGGATGATGAGCGCTGCTGATCAGGATGAGCGTCTCCCTGAGC	1984
Db	1095	AGTTCTCCCTGGAGGGACATGATGACCTTGTGTGATGAAAAAGACACTCTGACAGAAG	1154
QY	1985	ACCTGGCCGCTTCTACATTCGACAGTGTGACCCCTGGCCATTGAAAGTGTCCACAAAGATG	2044
Db	1155	AGGAGACTCACTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTCAACCACTG	1214
QY	2045	GCTTTATCCACCGGGACATCAAGCCTTGACANCACTACTCATCGACCTGGATGGTCAATATTA	2104
Db	1215	GATTCATCCACAGACATCAACACAGACACTCTCTTTTGGACAGCAAGGCCCATGTGA	1274
QY	2105	AGCTGACAGATTTTGGCCTCTGCACCTGGATTTCAGGTGGACTCAAAATTCCAAGTACTACC	2164
Db	1275	AACTTTCTGACTTTGGTCTTTGACAGGACTCGAAAAAGCACATAGACAGCAATTTTATA	1334
QY	2165	AGAAAGGACCAACATGACAGACAGACATGGAGCCCGGTGACCTCTGGGACGATGTTT	2224
Db	1335	GGAATCTGAACACAGCCTCC-----CCAGTGATTTTCAGTTCTCCAGAACATCAAT	1384
QY	2225	CCAACTCTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCGCAGAACGACGACC	2284
Db	1385	TCCAAAAGGAAAGCAGAAACCTGGAAAAGAAATAGACGTGAG-----1426	
QY	2285	AGAGTGCCCTGGCACATTTCTCTGTGCGGACACCAAAATACATCGCTCCGGAGGTGCTTC	2344
Db	1427	-----CTAGCCTTCTCCACAGTAGGCATCTCTGACTTCCTGAGCTACATTTCTCTGAGGTCTTCA	1478
QY	2345	TCCGAAAGGTTACACGACGCTCTGTGACTGTGTGGAGCGTCGGTGTGATTTCTTCTTTGAGA	2404
Db	1479	TGCAGACCGGGTACACAAGCTCTGTGATTTGGTGGCTGGGGTGATCATGTATGAGA	1538
QY	2405	TGCTGGTTGGGACGCGCCTTTCTTGGCCCCCACCACACAGAGACGAGCTGAAGGTGA	2464
Db	1539	TGCTCATGCGCTACCCACCTTTCTCTTCGAGACCCCTCAAGAGACATATAAGAGGTGA	1598
QY	2465	TCAACTGGGAGACGCTGTGCATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCCCGAG	2524
Db	1599	TGAACGGAAAGAACTTTGACTTTTCTCCAGAGTTTCCATCTCTGAAAGCCAGG	1658
QY	2525	ACCTCATCAGAGCTGTGCTGCGGCGCTGACTGCCCGCTGGGCGAGGATGGGGCAGATG	2584
Db	1659	ATCTAATTTTGAAGTTCTGCTGTGAATGGGAACATAGAATTTGGAGCTCTCTGGATTGAGG	1718
QY	2585	ACCTCAAGGCACACCGTTCTTCAACACCATCTGACTTTCCCGTGACATCCGAAGCAGG	2644
Db	1719	AAATAAAAGTAACCTCTTTTGAAGCGGTTGACTGGGAACATATCGAGAGAGACCTG	1778
QY	2645	CTGCACCTTACGTCCCCACCATCAGC	2670
Db	1779	CTGCATATCTATTGAAATCAAAAGC	1804

EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05052
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: 94810746.1
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 3018
TYPE: DNA
ORGANISM: Homo sapiens.
FEATURE:
NAME/KEY: CDS
LOCATION: (596)..(1990)
US-09-338-132-6

Query Match 7.1%; Score 225.2; DB 3; Length 3018;
Best Local Similarity 52.2%; Pred. No. 1.7e-49;
Matches 577; Conservative 0; Mismatches 493; Indels 36; Gaps 2;
QY 1565 AGCTGGACGAGGAATGCCAAAGCTGGCTCTGTGAGCGCGAGCAGGAGGATGAGGA 1624
DB 735 AGTTAGAAAGGTGATGAAGAGAGAGGCGCTTAAAGATGAGGAGAAACGACTCCGGAGAT 794
QY 1625 AGATCCTCTACCAAGAGGAGTCTAACTACAAACCGCTGAAGAGGCGCAAGATGGACAAGT 1684
DB 795 CAGCACATGCTCGGAAGAAACAGAGTTCTTCGTTTGAAGAGAAACAGACTTGGATTGG 854
QY 1685 CAYGTTTGTGAATCAAGACTCTAGGCATCGGTGGCTTTGGGGAAGTGTGCCCTTCGTT 1744
DB 855 AAGATTTTGTAGTCCCTTAAAGAGTAATPAGGCAGAGGAGCATTTGGTGAGGTACGCTTGTTC 914
QY 1745 GTAAGCTGAGCACTCACGCTCTGTAGCGCATGAGCACTCTCAGGAAGAGGATGCTCTGA 1804
DB 915 AGAAGAAAGATACGGGACATGTGTGCAATGAAATTAATCTCGTAAACAGATATGCTTG 974
QY 1805 ACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGAGCATCTCTGCTGAAGCAGACAATG 1864
DB 975 AAAAGAGCAGGTTGGCCACATTCGTGCGGAGCGTGACATCTAGTGGAGGCGAGCAGATT 1034
QY 1865 AGTGGTGTCAACTCTACTACTCTCTCCAGGACAAGCAGCCCTGTACTTTGTGATGG 1924
DB 1035 TGTGGGTTGTGAAATGTCTATAGTTTTCAGGATTAAGCTAAACCTCTACCTTAATCATGG 1094
QY 1925 ACTACATACCAGCGGGATGATGATGAGCTGTCTCAGGATGAGGTCTTCCCTGAGC 1984
DB 1095 AGTTCTGCTGGAGGAGCATGATGACCTGTGTGATGAAAGAACACACTCTGACAGAG 1154
QY 1985 ACCTGGCCCGTCTTACATTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAGATGG 2044
DB 1155 AGGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTCCACCACTG 1214
QY 2045 GCTTTATCCCGGACATCAAGCTTGACACATACTCATCGACCTGGATGTCATATTA 2104
DB 1215 GATTTCATCCACAGACATCAACACAGAACCTTCTTTTGGACAGCAAGGCGCCATGTA 1274
QY 2105 AGCTGACAGATTTTGGCTCTGCACTGGATTGAGCTGAGCTCACAAATCCAGTACTACC 2164
DB 1275 AACTTCTGACTTTGGCTTTGACAGAGCTGAAAAAGCAGCATAGGACAGATTTTATA 1334
QY 2165 AGAAGGGAACACATGACAGAGGACAGCATGGAGCCCGGTGACCTCTGGGACGATGTT 2224
DB 1335 GGAATCTGAACACAGCCTCC-----CCAGTGATTTCACITTTCCAGAACATGAAT 1384
QY 2225 CCAACTGTGCTGTGGAGCAGGTTAAAGACCTCTGGACAGAGGCGGCAAGAGCAGCACC 2284
DB 1385 TCCAAAGGAAAGCAGAAACCTGGAAAAAGAAATAGACGTCAG-----1426
QY 2285 AGAGTGTGCTGACATTTCTTTGCGGACACCAAAATTAACATCGCTCCGGAGGTCTTC 2344
DB 1427 -----CTAGCCTTCTCCACAGTAGGACATCTCTGACTACATTCCTCTGAGGTGTTCA 1478
QY 2345 TCCGAAAGGGTACACGAGCTCTGTGACTGTGTGAGGCGGTGCTGTGATTTCTTTTGA 2404

DB 1479 TGCACACCGGTTACAAAGCTCTGTGATTTGGTTCGCTTGGGTGATCATGTATGAGA 1538
QY 2405 TGCTGTTGGGAGCGGCTTCTTTGGCCCCCACCACACAGAGAGGAGCTGAAGGTGA 2464
DB 1539 TGCTCATCGGCTACCCACCTTTCTGTCTGAGACCCCTCAAGAGACATATAAGAAGGTGA 1598
QY 2465 TCAACTGGGAGAGCACGCTGTCATATCCCTACGACGTGAGGCTCAGCGCTGAGGCCGCG 2524
DB 1599 TGAAGTGAAGAAACTTTTGACTTTTCTCCAGAAGTTCCCACTCTCTGAGAAGCCAAAG 1658
QY 2525 ACCTCATCAGGAAGCTGTCTGCGCGGCTGACTGCCCTGGCGAGGATGGGCGAGATG 2584
DB 1659 ATCTAATTTTGAAGTTCTGCTGTAATGGGAACATAGAATTTGAGCTCTCTGAGTTGAGG 1718
QY 2585 ACCTCAAGCAGCACCGCTTCTTCAACACCATGACATTTTCCCGTGCATCCCAAGACAGG 2644
DB 1719 AAATAAAAGTAACATCTTTTGTGAAGCGGTTGACTGGGAACATATCAGAGAGACCTG 1778
QY 2645 CTGACCCCTACGTCCTCCCAACATCAGC 2670
DB 1779 CTGCAATATCTATTGAATCAAAAGC 1804
RESULT 13
US-08-860-150-1
Sequence 1, Application US/088601508
Patent No. 5981205
GENERAL INFORMATION:
APPLICANT: Hemmings, Brian A.
APPLICANT: Millward, Thomas A.
TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
FILE REFERENCE: 4-20265/A/PCT
CURRENT APPLICATION NUMBER: US/08/860,150B
CURRENT FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05052
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: 94810746.1
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 2101
TYPE: DNA
ORGANISM: Drosophila melanogaster
FEATURE:
NAME/KEY: CDS
LOCATION: (132)..(1499)
US-08-860-150-1

Query Match 7.1%; Score 222.6; DB 2; Length 2101;
Best Local Similarity 52.5%; Pred. No. 6.9e-49;
Matches 557; Conservative 0; Mismatches 474; Indels 30; Gaps 2;

QY 1563 ACAGCTGGAGCAGGAAATGGCCAAAGCTGCTCTGTGAGGCGCGAGCAGGAGCATGAG 1622
DB 281 AAAGCTGGAGGCTCAGCTCAAGGACGAGAGCTTGTGCGAGGCGCAGCGCCAGGAAGCG 340
QY 1623 GAAGATCTCTTACCAAGAGGAGTCTAATCAACCGGTGAAGAGGGCCCAAGATGGACAA 1682
DB 341 TCTGACGATGCCCAAGAGGAGCAGGATATCTCCGGCTGAAGCGATTTCGCGCTCGGTGT 400
QY 1683 GPCCATGTTGTGAAATCAAGACTCTAGGCATCGTGGGAGTCTGCTGCTGCTGCTGCTG 1742
DB 401 GGAGACTTTTGAAGCCCTCAAGTCAATCGAGCGCGCGGCTTCGGTGAAGTTCGGTTGT 460
QY 1743 TTGTAAGCTGGACACTCAGCTCTCTAGCCCATGAAGACTCTCTCAGGAAGAGGATGCTCT 1802
DB 461 GCAGAAAAGGACACTGGACATGTCTGCCCATGAAGTGTCTGCCCAAGCGACATGCT 520
QY 1803 GAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAG---GGACATCTCGGCTGAAGCAGA 1859
DB 521 GGAAGAGGAGGAGGTGGCACACGCTACGGCGCCGAGGCTGTGATGTCTTCTGTCGAGGCCGA 580

QY 2460 GTGATCAACTGGAGAGACGCTGCATATCCCTAGCAGGTGAGGCTCAGCGCTGAGGC 2519
 Db 1154 GTGATGAAGTGGCGGAGACGCTGATATTTCCCCCAAGAGATCCCATATCGAGGAGGC 1213
 QY 2520 CCGAGACCTATCAACGAAGCTGTGTCGCCGCTGACTCCGCGCTGGGAGGATGGGC 2579
 Db 1214 CAAGGAGAGCATCATCAACTTCTGTGCGAGGCGGATCGCGCTGGTTCCAGCGTCGTC 1273
 QY 2580 AGATGACCTCAAGGCACACCCGCTTCTCAACACCATCGACT 2620
 Db 1274 TGAGAGATCTGAAGTCGTCGCTTCTTCCGGGAGTTGACT 1314

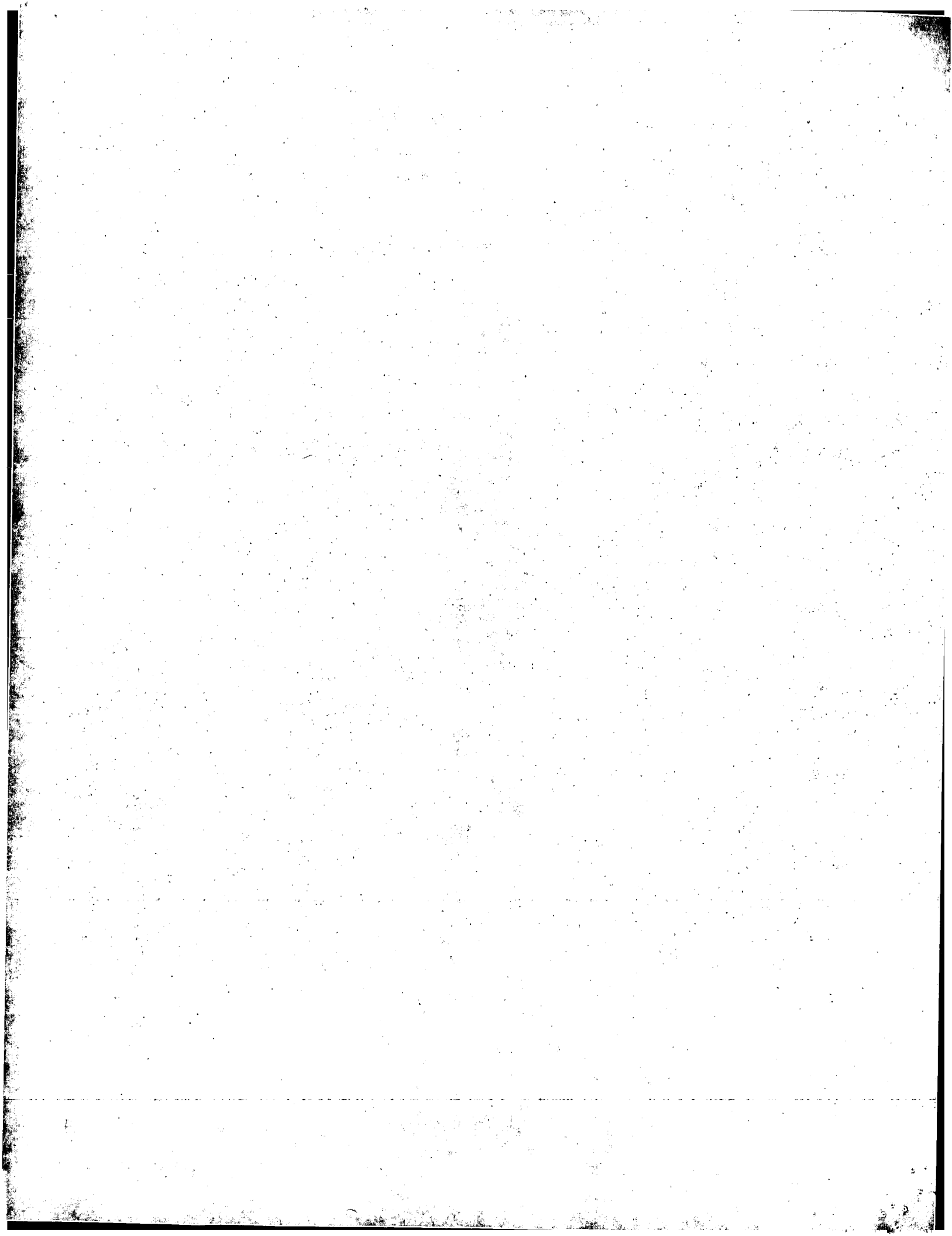
RESULT 15

US-09-328-111-66
 ; Sequence 66, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; CURRENT FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; EARLIER FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows version 3.0
 ; SEQ ID NO 66
 ; LENGTH: 678
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(678)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-328-111-66

Query Match 6.8%; Score 214.4; DB 4; Length 678;
 Best Local Similarity 68.8%; Pred. No. 6e-47; Mismatches 138; Indels 2; Gaps 1;
 Matches 308; Conservative 0;
 QY 1529 TCAAAACCTTACCAGCAGAGAGGTTCAGCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAG 1588
 Db 3 TCAAAATCTCATCAGCAGCGCTTACATCGTAAAAACAATTAGAGATGAATGATCCGGG 62
 QY 1589 CTGGGCTCTGTGAGGCGGAGCAGGAGCATGAGGAAGATCTCTACCAAGAGGAGTCTTA 1648
 Db 63 TTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAAGATCTA 122
 QY 1649 ACTACAACCGGCTGAGAGGCGCAAGATGGACAAGTCCATGTTTGTGAAAAATCAAGACTC 1708
 Db 123 ATTACATCCGCTTAAAGGGGCTAAATGGACAAGTCTATGTTTGTGAAGATAAAGACAC 182
 QY 1709 TAGGCATCGCTTTGGGGAAGTGTGCTCGCTTGTAAAGCTGGACACTCAGCTCTGT 1768
 Db 183 TAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAAAAAGTAGATACTAAGGCTTGT 242
 QY 1769 ACGCCATGAAGACTCTCAGGAAGAAGAGTGTCTTGAACCGGAATCAAGTGGCCCATGTCA 1828
 Db 243 ATGCAAGAAAAAATCTTCGAAAGAAGAGTGTCTTCTTCTCGAAATCAAGTCGCTCATGTTA 302

QY 1829 AGGCTGAGAGGAGACATCCCTGGCTGAAGCAGACAATCAGTGGTGGTCAAACTCTACT 1888
 Db 303 AGGCTGAGAGAGATATCTCTGGCTGAAGCTGACAATGAATGGTAGTTCCTCTATATT 362
 QY 1889 CTTCCAGGACAAGGACAGCCTGTACTTTGTGATGGACTACATACACGCGGGGATATGA 1948
 Db 363 CATTCCAAGATAAGGCG--CATTTATCCTTGTAAATGGCTACATTCCTNGGGGTGATATGA 420
 QY 1949 TGAGCCTGCTGATCAGGATGGAGGCTTT 1976
 Db 421 AGAGCCCATTAATANAATGGCATCTT 448

Search completed: January 16, 2003, 21:57:26
 Job time : 134.191 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:42 ; Search time 87.5516 Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagccaccgcgaagt.....aagagcattatttggggg 3155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 33868 seqs, 22934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	1035.2	32.8	2043	9	US-09-836-392-2
2	546.6	17.3	676	9	US-09-764-868-214
3	333	10.6	638	10	US-09-879-536-26
4	231.2	7.3	1689	9	US-09-938-842A-1861
5	228	7.2	1452	9	US-09-938-842A-2402
6	225.2	7.1	3583	9	US-09-974-298-152
7	214.4	6.8	678	10	US-09-879-536-66
8	192.2	6.1	734	9	US-09-764-868-196
9	169.4	5.4	1818	10	US-09-771-161A-89
10	161.8	5.1	2549	10	US-09-880-107-3691
11	155.6	4.9	1635	10	US-09-880-107-2340
12	149.8	4.7	3407	10	US-09-971-845-1
13	148	4.5	1398	9	US-09-938-842A-633
14	138.8	4.4	1416	9	US-09-938-842A-2503
15	135.6	4.3	1735	9	US-09-764-868-58
16	134.8	4.3	1244	10	US-09-771-161A-38
17	134.8	4.3	1393	10	US-09-771-161A-37
18	133.6	4.2	1257	10	US-09-799-875-15
19	133.6	4.2	1826	10	US-09-799-875-13

20	133.6	4.2	1972	9	US-10-098-841-166	Sequence 166, App
21	129.2	4.1	3061	10	US-09-880-107-2146	Sequence 2146, App
22	126.4	4.0	2610	10	US-09-970-000-3	Sequence 3, Appli
23	125.8	4.0	1515	10	US-09-804-471A-1	Sequence 1, Appli
24	125.8	4.0	5877	12	US-10-028-946-3	Sequence 3, Appli
25	125.8	4.0	6165	12	US-10-028-946-1	Sequence 1, Appli
26	122.6	3.9	6159	9	US-10-017-216-3	Sequence 3, Appli
27	122.6	3.9	6574	9	US-10-017-216-1	Sequence 1, Appli
28	120.2	3.8	568	9	US-09-796-692-8122	Sequence 8122, App
29	120.2	3.8	6409	10	US-09-864-864-293	Sequence 293, App
30	119.8	3.8	277	10	US-09-294-093B-2043	Sequence 2043, App
31	118.4	3.8	3244	9	US-10-174-590-571	Sequence 571, App
32	118.4	3.8	3244	9	US-10-176-758-571	Sequence 571, App
33	118.4	3.8	3244	12	US-10-052-586-571	Sequence 571, App
34	118.2	3.7	3611	9	US-09-836-392-3	Sequence 3, Appli
35	118	3.7	362	9	US-09-796-692-7789	Sequence 7789, App
36	111.8	3.5	2146	10	US-09-954-456-527	Sequence 527, App
37	111.8	3.5	2146	10	US-09-842-307-1	Sequence 1, Appli
38	108.4	3.4	1479	10	US-09-771-161A-46	Sequence 46, Appli
39	106.6	3.4	512	10	US-09-867-701-6230	Sequence 6230, App
40	105.8	3.4	5313	10	US-09-801-368-297	Sequence 297, App
41	103.4	3.3	678	10	US-09-841-683-4	Sequence 4, Appli
42	103.4	3.3	711	10	US-09-841-683-6	Sequence 6, Appli
43	103.4	3.3	1191	10	US-09-841-683-10	Sequence 10, Appli
44	103.4	3.3	1224	10	US-09-841-683-8	Sequence 8, Appli
45	103.4	3.3	1485	10	US-09-801-876B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; TITLE OF INVENTION: Antibodies
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match 32.8%; Score 1035.2; DB 9; Length 2043;
Best Local Similarity 85.4%; Pred. No. 8.2e-271;
Matches 1170; Conservative 0; Mismatches 188; Indels 12; Gaps 1;

QY	1585	AAAGCTGGGCTCTGTGAGCGGCGAGCAGGAGGAGATGAGAGATCCTCTACCAAGAGAG 1644
DB	215	AATGCTGGACCTCTGTGAGCTGAGCAGGAGGAGATGCGGAGAGATCCTCTACCAAGAGAG 274
QY	1645	TCTACTACTACACCGCTCAAGAGGCGGCGAGGAGGAGATGAGAGATCCTCTCTGAAATCAAG 1704
DB	275	TCTATTACACAGCTGTAAGAGGCGGCGAGGAGATGAGAGATCCTCTCTGAAATCAAG 334
QY	1705	ACTCTAGGATCGGGTGGCTTTTGGGGAAGTGTGCTTCGCTTGTGAAGCTGACACTCACGCT 1764
DB	335	ACCCTGGGGAGCGTGGCTTTTGGGGAAGTGTGCTTCGCTTGTGAAGCTGACACTCACGCT 394

[illegible]

```
Db 539 GTCCACAAGATGGCTTATCCACCAGACATCAAGCCTCATAAACATTTTGATAGATCTG 598
QY 2092 GATGGTCATATTAAGCTGACAGATTTTGGCTCTGCTGACATGGATTCAGGTGGACTCACAAT 2151
Db 599 GATGGTCATATTAAGCTGACAGATTTTGGCTCTGCTGACATGGATTCAGGTGGACTCACAAT 658
QY 2152 TCCAAGTACTACCAGAAA 2169
Db 659 TTCCAATATTACCAGAAA 676

RESULT 3
US-09-879-536-26/c
; Sequence 26, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 26
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-536-26

Query Match 10.6%; Score 333; DB 10; Length 638;
Best Local Similarity 71.3%; Pred. No. 1.6e-80;
Matches 454; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1525 GTCATCAAACTACCAGCAGAGGTTCAGCGGAGGCTACAGCTGGAGCAGGAATGGCC 1584
Db 637 GTACTCAAACTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAGAATGAATGATG 578
QY 1585 AAGCTGGGCTGTGTAGGCCGAGGAGGAGATGAGGAAGATCTCTACAGAAAGGAG 1644
Db 577 CGGGTTGGATTATCTCAAGATCCCGAGGATCAATGAGAAAGATGCTTTGCCAAAGAA 518
QY 1645 TCTAATAACAACCGCTGAAGAGGCCAAGATGGACAAGTCCATGCTTTGTGAATCAAG 1704
Db 517 TCTAATATCATCCGCTTAAAAGGGCTAAAATGGACAAGTCTGATGTTTGAAGATAAG 458
QY 1705 ACTTAGGCTCGGTGCCCTTTTGGGAAGTGTGCCCTGCTTTGAAGTGGACACTCACGCT 1764
Db 457 ACACATAGGAATAGGACATTTGGTGAAGTCTCTAGCAAGAAAGATAGATACTAAGCT 398
QY 1765 CTGTAGGCCATGAAGACTCTCAGGAAGAGATGCTTGAACCCGGAATCAAGTGGCCCAT 1824
Db 397 TTGTATGCAACAAAAACTCTTCGAAATAAAGATGTTCTTCTCGAAATCAAGTCGCTCAT 338
QY 1825 GTCAAGGCTGAGGGACATCTCTGCTGAGCAGACAAATGAGTGGTGGTCAACTCTAC 1884
Db 337 GTTAGGCTGAGAGATATCTTGCTGAAGCTGACAATGAATGAATGAGTGGTACTGCTATAT 278
QY 1885 TACTCCTCCAGGACAAGGACAGCCTGTACTTTGTGTGATGAGTACTACATACCAGCGGGAT 1944
```

```
Db 277 TATTCATTCCAAGATAAGCACAATTTATATCTTTGTAAATGACTACATTCCTGGGGTGAT 218
QY 1945 ATGATGAGCCTGCTGATCAGGATGGAGTCTTCCCTGAGCACCCTGCCCGCTTCTACATT 2004
Db 217 ATGATGAGCCTATTAATTAAGATGGGCATCTTTCCAGAAAGTCTGGCAGCATTCACATA 158
QY 2005 GCAGAGTTGACCTGGCCATTTGAAAAGTG--TCCACAAGATGGCTTTTATCCACCGGAC 2061
Db 157 GCAGAACTTACCTGTGTCAGTTTGAAGGTGTTCCCTTAAATGGGTTTATTCATAGAT 98
QY 2062 ATCAAGCCTGACAACATACATCTGACCTGGATGTCATATTAAAGTCAGAGATTTTGGC 2121
Db 97 ATTAACCTGATAAATATTTTGAATGATGTCATATTAAATGACTGACITTTGCG 38
QY 2122 CTCTGCATGGATTCAGGTGGACTCAAAATTCCAAGT 2158
Db 37 CTCTGCATGGCTTTCAGATGGACACAGATTCCTAAGT 1

RESULT 4
US-09-938-842A-1861
; Sequence 1861, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS, CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1861
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1861

Query Match 7.3%; Score 231.2; DB 9; Length 1689;
Best Local Similarity 51.8%; Pred. No. 1.2e-52;
Matches 580; Conservative 0; Mismatches 528; Indels 12; Gaps 2;

QY 1504 ATGGAGCAACACGTGGAGAAATGTCATCAAACTTACCAGCAGAAAGTCAAGCGGAGGCTA 1563
Db 172 ATCGAATCATTTATAAAATTCAGAAAGAGAGTCTCCAGAAAGAAAGAGCGTCGGAGC 231
QY 1564 CAGCTGGAGCAGAAATGGCCAAAGCTGGGCTCTGTGAGCGCCGAGCAGAGCAGATGAGG 1623
Db 232 ATCTTGGAAACAACCTAGCTGATGCTGATGTTTGAAGACAAGATGATATATTA 291
QY 1624 AAGATCCTCTACAGAAGAGGTCTAACTACAACCCGGCTCAAGAGGGCCCAAGATGGACAAG 1683
Db 292 AAGAATTTGACAAAAGGAATGAGTATATGCGTCTACAAAGACAGAAATGGGGTT 351
QY 1684 TCCATGTTTGTGAATAATCAAGACTCTAGGCATCGGTGCGCTTTGGGGAAGTGTGCTCGCT 1743
Db 352 GATGACTTTGAATGCTGTAGCATCATTTGCCGGGGTGTCTTCGGTGAAGATTTCT 411
QY 1744 TGTAGCTGGACACTCACGCTCTGTACCCATGAAGACTCTCAGGAGAGAGATGTCCTG 1803
Db 412 AAAGAAAAATCTACTGGGAAGCGCTATATGAATGAAGAAAGTTAAAGAAATCCGAGATGCT 471
QY 1804 AACCCGAATCAAGTGGCCCATGTCAGAGGTGAGAGGAGACATCTGGCTGAAGCAGACAAT 1863
```

Db 472 CGACGAGGCGAGTGGACATGTTAAAGCTGAAGAAATGCTTGAGAGTGGATAGC 531
QY 1864 GAGTGGGTTGCAAACTCTACTACTCTCCAGGACAGGACACCTGTACTTTGTGATG 1923
Db 532 CTTTCATTTGCAAGCTTTGTTACTCTTCCAGATGATGACATTTGTATCTATTATG 591
QY 1924 GACTACATACAGGCGGGGATATGATGAGCTGCTGTATGAGTGGAGGCTTCCCTGAG 1983
Db 592 GAATACCTCCCTGGAGGTGATGATGACACTGCTGATGCGAAGGATACCTTACGGGAA 651
QY 1984 CACCTGGCCCTCTCATATGACAGAGTGGACCTGCGCATTTGAAAGTGTCCACAAGATG 2043
Db 652 GATGAGACTCGTTTATGTTGGCGACACAAATTTGCTATGAGTCTATCCATAAGCAT 711
QY 2044 GGCTTTATCCACGGGACATCAAGCTTCACAACATCTCATCCACCTGGATGTCATATT 2103
Db 712 AATTACGTTCCACAGGATATAAGCCCTGATATTTATGTTACTCGAAACGGCCATATC 771
QY 2104 AAGCTGACATTTTGGCTCTGCATGAGTTCAGTGGACTCACAAATTTCCAGATTTTAAG 831
Db 772 AAGCTTTGAGATTTGGATTGAGCAAGTCTCTGGAAGCAAAAATTTCCAGATTTTAAG 831
QY 2164 CAGAAGGGAACACATGAGACAGGACATGAGCGCCGCTGAGCAGAGGCGCA 2223
Db 832 CGCGAGCTTTGTCACAGG-----AGTACAAAGCCTGCAGAGAACATGATAGACTC 882
QY 2224 TCCAACCTGCTGTGGAGACAGGTTTAAAGACCCCTGGAGCAGAGGCGCA 2280
Db 883 TCCAAGCTCTCTGACCTAGGAGAACTCAGCAGGAAACAGCTTTTACATTTGGCAACA 942
QY 2281 CACGAGAGTGGCTGGACATCTCTTTCGCGGACACCAATATACATCGTCCGAGGTG 2340
Db 943 AACAGAAGGACCTGCTTTCTACAGTAGGAATCCCGATTAATGCTGAGTCTGATGTC 2400
QY 2341 CTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGGCTGCGGTGATGTTCTTT 2400
Db 1003 CTGCTGAAGAAAGGTTATGGAATGAGTGTGATTGTTGCTCTTGGAGCAATCATGTC 1062
QY 2401 GAGATGCTGTTGGCAGCGCTTCTTTCGCGGACACCAATATACATCGTCCGAGGTG 2460
Db 1063 GAGATGCTGAGGTTTTCGCGGCTTCTTTCGCGGACACCAATATACATCGTCCGAGGTG 1122
QY 2461 GTGATCACTGGAGAGCAGCTGATATCCCTACGAGGTGAGGCTGAGGCTGAGGCC 2520
Db 1123 ATGTAACTTGAACACCTGCTGAAGTTCCTGATGAGCTAAGCTCTCCATCGAGGTA 1182
QY 2521 CGAGACCTCATCAGAACTGCTGCTGCGGCTGACTGCGCGCTGAGGAGGATGGGCA 2580
Db 1183 AAGATCTTATCCGAAGACTCTCTGCAATGTGCAACAGAGGCTTGGGACCAAGAGTT 1242
QY 2581 GATGACCTCAAGGACACCCCTTCTTCAACACCATGCACT 2620
Db 1243 CACGAAATCAAGACACACCTTGGTTAGGGGAGTCAAT 1282

RESULT 5
US-09-938-842A-2402
; Sequence 2402, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krepis, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09938842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16

Query Match
Best Local Similarity 7.2%; Score 228; DB 9; Length 1452;
Matches 638; Conservative 0; Mismatches 640; Indels 6; Gaps 2;
QY 1433 AGAATAGCAGAGATGAAGAGAGAGAGTCTCCATCAAGAGTTACTCCCTTATGCTT 1492
Db 23 AGAAGGTACGAGAGAGGTTCTGGATCGAGCTTGACCATGGAGAAAGTGGCCGAG 82
QY 1493 TCAATTTCTTCAATGAGCAACACGTTGGAGATGTCTCAAAACCTACACAGAGGTCA 1552
Db 83 CGAAGCAGTACATCGAATCACTACAAAGCTCAGAACAGACATTCAGAGAGGAAG 142
QY 1553 GCCGAGGCTACAGCTGAGCAGGAAATGSCAAAGCTGGGCTCTGTGAGGCGCAGAG 1612
Db 143 AGACACGCTGATCTTGAAGAAAGTTGGCTCTTCTGGAGTGCACAAAGAGAGCAAA 202
QY 1613 AGCATGAGGAGATCTCTTACGAGAGGAGTCTAACTACACCGCTGAGAGAGGCA 1672
Db 203 TCAACATGATTAAAGACCTGGAGAGAAAGAGACAGAGTTTATGAGCTTTAAAGGA 262
QY 1673 AGATGGACAAATCCATCTTTGTCAAAATCAAGACTTAGGCATCGCTTTGGGGAAG 1732
Db 263 AGATAGTGTGATGACTTTTGGAGCTTTTGAATATTTATGCCATGAGAGTTAAAGAA 382
QY 1733 TGTGCTCTGCTTAAAGCTGGACACTCAGCTCTGTACCCCATGAGACTCTCAGAGA 1792
Db 323 TTGCTCTATGCTGGAGAGAAAGTCTGGAATATTTATGCCATGAGAGTTAAAGAA 382
QY 1793 AGGATCTCTGACCGGATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCTGGCTG 1852
Db 383 CTGAATGCTGATGAGAGAGAGGTTGAGCATGTGAGAGAGAGGAACTGCTGGCTG 442
QY 1853 AAGCAGCAATGAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGACACCTGCT 1912
Db 443 AGGTTGAAAGCCATATATTTGAGAGCTTACTATTTCAITTCAGGATCCCGAGTATCT 502
QY 1913 ACTTTGTGATGACTACATACAGGCGGGGATGATGATGAGCTCTGATCAGGATGAGG 1972
Db 503 ATCTGATTATGAAATATCTCCCGGTGATGATGATGATGATGATGATGATGATGATG 562
QY 1973 TCTTCCCTGAGCACCTGCGCTCTTACATTTGAGAGTGTGACCTTGGCCATTTAAAGTG 2032
Db 563 CATTACGGGAGATGTTGCCAGATTTTATTTGCTCAAGTGTCTTGGCTTGAATTTAC 622
QY 2033 TCCACAAGATGGCTTTATCCACCGGACATCAAGCTTGACCAACATCTATCTGACCTG 2092
Db 623 TACACAGTACAACTATATTTATAGGAGATCAAACTGATACCTTCTTTGGACAAAG 682
QY 2093 ATGTCATATTAAAGTACAGATTTTGGCTCTGCACTGATGATGATGATGATGATGAT 2152
Db 683 ATGGCAGATCAAACTCTCGGACTTTGGGCTTGTAAAGCTCTTGTGATTTAGAAATTTAC 742
QY 2153 CCAAGTACTACGAAAGGAAACACATGAGACAGGACAGCATGGAGCC---CGGTGACC 2209
Db 743 CTTCATTTACGAGGAAATGAGCCAGCATGATGAAACTATGTCAGAACCTATGATGATG 802
QY 2210 TCTGGGACGATGTTTCCAACTGTGCTGTGGACAGGTTTAAAGACCTGGAGCAGAGG 2269
Db 803 ATAGATCTTCTGACACTGATAAAGAGAGCTGGCGCAGTCCCGCAGCAAACTTC 862
QY 2270 CGCAGAGCA---GCACAGAGGTGCTGACATTTCTTGTGCGGACACCAAAATTACA 2326
Db 863 AGCATTTGGCAGATGAATCGCAGAAACTAGCATTTTCACTGTGGAAACACCGGACTATA 922

QY 2327 TCGCTCCGGAGGTGCTTCTCCCAAGGGTACACGAGCTCTGTGACTGGTGGAGCGTGG 2386
 Db 923 TTGCTCCTGAAGTCTTGTGCTGAAGAAAGGATATGGCATGGAATGTGATTGGTGTCTATTAG 982
 QY 2387 GTGTGATCTCTTTGAGATGCTGGTGGCCAGCCGCTTTCTTTGGCCCCCAGCCAGAG 2446
 Db 983 GTGCAATATGATGACAAATGCTGCTGGGTATGCTTCTTCTTTATGCTGATGACCTATAT 1042
 QY 2447 AGACGAGCTGAAGGTGATCAACTGGGAGAGCACGCTGCATATCCCTACGCGAGGTGAGGC 2506
 Db 1043 CAACITGCGAAGATGCTCAATGGGAAACCAATTTGAAATTTCTGAGGATGGAAGT 1102
 QY 2507 TCAGCGCTGAGCCCGAGACCTCATCAGAAAGCTGTGCTGCGGCTGACTCCGCCCTGG 2566
 Db 1103 TTTTCATCTGAGCAAAAGATCTTATGTCAGGTGCTGTGCAACGTTGACCATAGGCTTG 1162
 QY 2567 GCAGGATGGGAGACCTCAAGGCACACCCGTTCTTCAACACCATCGACTTTTCCC 2626
 Db 1163 GTACTGGAGGAGGCCCGAGCAATCAAGGATCATCTTGTTCAGAGGATGTTGTGGG 1222
 QY 2627 GTGACATCCGAAGCAGGCTGCACCCCTACGTCCCCACCATCAGCCACCCATGGACACCT 2686
 Db 1223 ARAAGCTCTATGAATGAGGCTGCGTCAACACAGAGTGAACGAGCTAGATACAC 1282
 QY 2687 CCAATTTGACCCGGTGGATGAAG 2710
 Db 1283 AAAATTTATGAATTTGATGAAG 1306

RESULT 6

US-09-974-298-152
 ; Sequence 152, Application US/09974298
 ; Patent No. US20020156263A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Chen, Hwei-Mei
 ; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
 ; FILE REFERENCE: PA-0037 P
 ; CURRENT APPLICATION NUMBER: US/09/974,298
 ; CURRENT FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: 60/238,331
 ; PRIOR FILING DATE: 2000-05-10
 ; NUMBER OF SEQ ID NOS: 194
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 152
 ; LENGTH: 3583
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20020156263A1 347975.11
 US-09-974-298-152

Query Match 7.1%; Score 225.2; DB 9; Length 3583;
 Best Local Similarity 52.2%; Pred. No. 8.3e-51;
 Matches 577; Conservative 0; Mismatches 493; Indels 36; Gaps 2;
 QY 1565 AGCTGAGCAGCAATGGCCAAAGCTGGCTCTGTGAGCCGCGAGGAGCAGATGAGGA 1624
 Db 423 AGTTAGAAAGGTGATGNAAGAAAGAGGCGCTAAAGATGAGGAGAAACGACTCCGAGAT 482
 QY 1625 AGATCCTCTACAGAAAGAGTCTAACTACACCGGCTGAAGAGGGCCAGATGGACAAGT 1684
 Db 483 CAGCACATGCTCGGAAGAAACAGAGATTCTTCGTTTGAAGAGAAACAGACTTGGATTGG 542
 QY 1685 CCATGTTTGTGAATCAAGACTCTAGGCATCGGTGCTTGGGGAAGTGTCCCTCGCTT 1744
 Db 543 AAGATTTTGTGCTGCTGAAAGTAAATAGGAGGAGGAGCATTTGGTGGAGTACGGCTTGTTC 602
 QY 1745 GTAAGCTGGACATCAGCGCTCTGTACGCGCATCAAGACTCTCAGGAAGAGGATGCTTCA 1804
 Db 603 AGAAGAGATACGGGACATGTGTATGTCATGAAATCTCCGTAAAGCAGATATGCTTG 662
 QY 1805 ACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGACATCTGCTGAGCAGACAATG 1864

Db 663 AAAAAGAGCAGGTTGGCCACATTCGTGCGGAGCGTGACATTCCTAGTGGAGGCGACAGTT 722
 QY 1865 AGTGGGTGGTCAAACTCTACTACTCTCTCCAGGAGCAAGAGACGCTGTACTTTGTGTATGG 1924
 Db 723 TGTGGTGTGTAAGAAATGTTCTATAGTTTTCAGAGTAAGCTAAACCTCTACCTAAATCATGG 782
 QY 1925 ACTACATACCGAGCGGGGATATGATGAGCTCTGATCAGGATGAGGCTTCTCCCTGAGC 1984
 Db 783 AGTTCTGCTGGAGGGGACATGATGACCTTGTGATGNAAGAACACACTCTGACAGAG 842
 QY 1985 ACCTGGCCGCTTCTACATTTGACAGTGTACCTGGCCATTTGAAAGTGTCCACAAGATGG 2044
 Db 843 AGGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTACCAACTTG 902
 QY 2045 GCTTTATCCAGCGGACATCAAGCTGACACATCTCATCAGCTGGATGCTCATATTA 2104
 Db 903 GATTCATCCACAGACATCAAAACAGACACCTCTTTTGGACAGCAAGGCCCATGTGA 962
 QY 2105 AGCTGACAGATTTTGGCCTCTGCACTGGATTGAGGTGGAATTCACAATTCAAAGTACTACC 2164
 Db 963 AACTTCTGACTTTGGTCTTTCACAGGACTGAAAGAACACATAGGACAGATTTTATA 1022
 QY 2165 AGAAGGGAACACATGAGACAGAGCATGGAGCCCGGTGACCTCTCTGGGAGCATGTTT 2224
 Db 1023 GGAATCTGAACACAGCCCTCC-----CCAGTGATTTCACCTTTCAGAACATGAAT 1072
 QY 2225 CCAACTGTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCGGAGAGCAGCACC 2284
 Db 1073 TCCAAAGGAAAGCAGAAACCTGGAAAGAAATAGACGTCAG----- 1114
 QY 2285 AGAGGTGCTGGCAGATTTCTTGTGCGGACACCAAAATTTACATCGCTCGGAGGTGCTTC 2344
 Db 1115 -----CTAGCCTTCTCCACAGTAGGACCTCTGACTACATCTGCTCTGAGGTGTTCA 1166
 QY 2345 TCCGAAAGGTPACAGCAGCTCTGTGACTGTGTGGAGCGTCTGGTGTGATTTCTTTTGA 2404
 Db 1167 TGCAGACCCGGTACAAAGCTCTGTGATTGGTGTGCTGGGTGATCATGTATGAGA 1226
 QY 2405 TGTCTGGTGGGCGCCGCTTTCTTTGGCCCCCAGCCACCCACAGACGAGCAGCTGAAGTGA 2464
 Db 1227 TGTCTATGGCTTACCCACCTTTCTTCTGTGAGACCCCTCAAGAGACATATAAGAAGTGA 1286
 QY 2465 TCAACTGGGAGAGCAGCTGTCATATATCCCTACGCGCTGAGGCTGAGGCGCGAG 2524
 Db 1287 TGAATGGAAGAAACTTTGACTTTTCTCCAGAGTTCCTCATCTCTGAGAAGCCAGG 1346
 QY 2525 ACCTCATCAGAAAGTGTGCTGCGGCGTGTACTGCGGCTGGGAGGAGTGGGAGATG 2584
 Db 1347 ATCTAATTTTGGAGTTCTGCTGTGAATGGAAACATAGAATTTGGAGCTCTCTGGAGTTGAGG 1406
 QY 2585 ACCTAAGGCGACACCGTCTTCTTCAACACATCGACTTTTCCCGTGACATCCCAAGCAGG 2644
 Db 1407 AAATAAAGTAACCTCTTTTTTTTGAAGCGGTGACTGGGAACATATCAGAGAGACCTTG 1466
 QY 2645 CTGCACTTACCTCCCGCCCATCAGC 2670
 Db 1467 CTGCAATATCTATTGAATCAAAAGC 1492

RESULT 7

US-09-879-536-66
 ; Sequence 66, Application US/09879536
 ; Patent No. US20020144298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan

Fri Jan 17 11:17:06 2003

APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/879,536
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/088,801
PRIOR FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 66
LENGTH: 678
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(678)
OTHER INFORMATION: n = A, T, C or G
US-09-879-536-66

Query Match 6.8%; Score 214.4; DB 10; Length 678;
Best Local Similarity 68.8%; Pred. No. 2.6e-48;
Matches 308; Conservative 0; Mismatches 138; Indels 2; Gaps 1;
QY 1529 TCAAAACCTACCAGCAGAGGCTACGCGAGGCTACAGCTGGAGCAGGAATGCCAAAG 1588
Db 3 TCAAAATCATCAGCAGCGCTACATCGTAAAAAACAATAGAGAAATGAATGCGGG 62
QY 1589 CTGGCTCTGTGAGCGCAGCAGGAGGATGAGGAGATCCCTACCAAGAGAGCTA 1648
Db 63 TTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAACTA 122
QY 1649 ACTCAACCGCTGAAGAGGCCAAGATGGACAAGTCCATGTTTGTGAAAATCAAGACT 1708
Db 123 ATTACATCCGCTTAAAGAGGCTAAATGGACAAGTCTATGTTGTGAAGATAAGACAC 182
QY 1709 TAGGATCGGTGCTTGGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1768
Db 183 TAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGATAGATTAAGGCTTGT 242
QY 1769 ACGCCATCAAGACTCTCAGGAAGAGGATGCTGCAACCGGAATCAAGTGGCCCATGCA 1828
Db 243 ATGCAACAAACACTTCCGAAGAAGATGTTCTTTCGAATCAAGTCTCATGTTA 302
QY 1829 AGGCTGAGAGGACATCTCGCTGAAGCAGACATGAGTGGGTGCTCAACTCTACTACT 1888
Db 303 AGGCTGAGAGAGATATCTCGCTGAAGCTGACAATGAATGAGTGTGCTATATTATT 362
QY 1889 CTTCCAGCAGCAGCAGCCTGTACTTGTGATGAGTACATACAGCGCGGGATATGA 1948
Db 363 CATTCAGATAAGGCG--CATTTATCTTGTATGGCTACATTCCTTNGGGGTATGA 420
QY 1949 TGAGCTCTGTATGATGAGGAGTCTT 1976
Db 421 AGAGCCCATTTATTAATGGGCATCTT 448

RESULT 8
US-09-764-868-196
Sequence 196, Application US/09764868
Patent No. US2002016871A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 196
LENGTH: 734
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-868-196

Query Match 6.1%; Score 192.2; DB 9; Length 734;
Best Local Similarity 56.2%; Pred. No. 2.9e-42;
Matches 424; Conservative 1; Mismatches 294; Indels 36; Gaps 2;

QY 1747 AAGCTGGACATCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGATCTCTCTGAC 1806
Db 12 AAGAAGATACAGGCCATATCTATCAATTAAGATATTGAGAAAGTCTGATGATGCTTGA 71
QY 1807 CGAATCAAGTGGCCCATCTCAAGGCTGAGGGGACATCTCGCTCAAGCAGCAATGAG 1866
Db 72 AAGAGCAGGTGGCCCATATCCGAGCAGAAAGATATTTGGTAAAGCAGCATGTTGCC 131
QY 1867 TGGGTGCTCAAACTCTTACTTCTCTCCAGGACAAAGCAGCTCTGTCTTGTGTGAGAC 1926
Db 132 TGGGTGCTGAAGATGTTTACAGTCTTTCAGGATAAGAGGAATCTTTATCTAATCATGGA 191
QY 1927 TACATACAGCGGGGATATGATGAGCTTGCATCAGATGAGGTCTTCCCTGAGCAC 1986
Db 192 TTTCTCCCTGGAGGTGACATGATGACATGCTTAATGAAGAAGACWCCCTGACAGAGAG 251
QY 1987 CTGGCCCGCTTCTACATTCAGAGTTCAGCCCTGGCCATTTGAAAGTGTCCCAAGATGGGC 2046
Db 252 GAACACAGTCTTACATTTTCAGAGACTGTTTGGCAATAGATGCGATCCACCAAGTTGGT 311
QY 2047 TTTATCCACCGGACATCAAGCTGACACATCTCATCGACCTGATGCTCATATTAAG 2106
Db 312 TTTATCCATCGGATATTAAAGCCAGACACCTTTTATGATGCCAAGGTCTATGTA 371
QY 2107 CTGACAGATTTTGGCTCTGCTGCTGATGATGAGTGCATCAATTTCCAAAGTACTACCAG 2166
Db 372 TTTATCTGATTTTGTGTTATGATCGGATTAAGAAGCTCACAGGACTGAATTTTATAGA 431
QY 2167 AAAGGGAACACATGACAGAGGACAGATGAGCCCGGCTGACCTCTGGGACGATGTTTC 2226
Db 432 AATCTCACACAC-----AACCCACCAAGTACTTCTCATTTTCAGAACATGAATC 481
QY 2227 AACTGCTGCTGTGAGACAGGTTAAAGACCTTGAGCAGAGGCGCAGAGCAGCAGCAG 2286
Db 482 AAAGAGAAAGAGCAAACTTG-----GAAGAAGACAGG 515
QY 2287 AGGTGCTGTCACATCTCTTGTGGGACACCAAAATACATCGCTCGGAGGTGCTTCTC 2346
Db 516 AGACAATGCGCATATCCACAGTTGGGACACAGATTTACATTTGCTCCAGAAATTCATG 575
QY 2347 CGCAAGGATACAGCAGCTCTGTGACTGTGGAGCGCTCGTGTGATCTCTTTGAGATG 2406
Db 576 CAGACTGTTTACACAAATTTGTGACTGTGTTGGGTGCTTTGGGAGTGAATGATGATGA 635
QY 2407 CTGTTGGGAGCGCGCTTTCTTGGCCCGCCACCCACAGAGAGCAGCAGCTGAAGGTGATC 2466
Db 636 CTAATAGATATCCACCTTCTGCTCTGAAACACTCAAGAGACATACAGAAAAGTGA 695
QY 2467 AACTGGGAGCAGCGCTGCTATATCCCTACGAGGT 2501
Db 696 AACTGGAAAGAACTCTGGTATTTTCTCCAGAGGT 730

RESULT 9
US-09-771-161A-89
Sequence 89, Application US/09771161A
Patent No. US2002010811A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26

```

; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-89

```

[illegible]

RESULT 10
US-09-880-107-3691
Sequence 3691, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Schafi, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression P

```

; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 3691
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No.
; NAME/KEY: unsure
; LOCATION: (1)..(2549)
; OTHER INFORMATION: n = a o r c o r g o r t
US-09-880-107-3691

```

Query Match	5.1%	Score 161.8	DB 10	Length 2549
Best Local Similarity	59.8%	Pred. No. 1.1e-33		
Matches 271	Conservative 0	Mismatches 182	Indels 0	Gaps 0
QY 1689	GTTTGTCAAAATCAAGACTCTAGGCATCGTGCCCTTTGGGGAAGTGTGCCCTCGCTTGTAA	1748		
DB 209	GTTTGAACGAATCAAGACCTCGCAGCGGCTCTCGGCGGGTGATGCTGGTCAACA	268		
QY 1749	GCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGGATGCTCTGAACCG	1808		
DB 269	CAAGGACCGCGGAACCACTATGCCATGAAGATCCTCGACAAAGAAGGTGGTGAACCT	328		
QY 1809	GAATCAAGTGGCCCATCTCAAGGCTGAGAGGACATCCTGGCTCAAGCAGACAATGAGTG	1868		
DB 329	GAACAGATCGACACACCCTGAATGAAGAGCGCATCTGCAAGCTGTCACTTCCGTT	388		
QY 1869	GGTGGTCAAACTTACTTACTCTTTCAGGACAAAGGACAGCCTGTACTTTGTATGGACTA	1928		
DB 389	CCTCGTCAAACTCGAGTTCCTTCAAGGACAACTCAAACTTATACATGGTCAATGGAGTA	448		
QY 1929	CATACAGGCGGGATATGATGAGCCTGCTCATCAGATGGAGTCTTCCCTGAGCACCT	1988		
DB 449	CGTGCCCGGGGGAGATGTTCTACACCTACGCGCGATCGGAAGTTTCAGTGAGCCCCA	508		
QY 1989	GGCCCGCTTCTACATTCGAGATTGACCTGGCCATTGAAAGTGTCCACAAGATGGCGTT	2048		
DB 509	TGCCCGTTTACCGGCCAGATCGTCCGTACCTTTGAGTATCTGCACCTCGCTGGATCT	568		
QY 2049	TATCCACCGGACATCAAGCCTGACACATCACTCACTGATCGATGGATGGTATTAAGCT	2108		
DB 569	CATCTACAGGACCTGAAGCGGAGATCTGCTCATTTGACCAAGCAGCGCTACATTCAGGT	628		
QY 2109	GACAGATTTGGCCTCTGCACTGGATTCAGGTG	2141		
DB 629	GACAGACTCGGTTTCCCAAGCGCGTGAAGGG	661		

RESULT 11
US-09-880-107-2340
; Sequence 2340, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherff, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054

Best Local Similarity 56.5%; Pred. No. 2,3e+20; Gaps 1;
Matches 299; Conservative

QY 1602 GGCCGACAGCAGATGAGGAAGATCCCTTACCAGAAGAGTGTAACATAACAACGGCT 1661
DB 899 GGCCCAAGCAAGTACGTGCCGACTTCCTTGCAGTGGCGGAGAGCCCATCGTTGGTAGGCT 958
QY 1662 GAAGAGGCCAAGATGACAAAGTCCATCTTTGTGAAAATCAAGACTCTAGGCATCGGTGC 1721
DB 959 TAAGAGGTCCGACTGACAGAGGACGACTTCGAGATTCTGAAGGTGATCGACACGCGGGC 1018
QY 1722 CTTTGGGGAAGTGTCCCTCGCTTTAAGCTGGACACTCACGCTCTGTACGCCATGAAGAC 1781
DB 1019 GTTCAGCGAGGTAGCGGTAGTGAAGATGAAGACAGACGGGCAGCGGTATGCCATGAAGAT 1078
QY 1782 TCTCAGGAAGAGATGTCTTGAACCGCAATCAAGTGGCCCCATCTCAAGCGCTGAGAGGGA 1841
DB 1079 CATGAACAGTGGGACATGCTGAAGAGGGCGAGGTGTCTGCTCCGTGAGGAGAGGGA 1138
QY 1842 CATCTGGCTGAAGCAGACAAATGATGGTGGTCAAACCTCTACTACTCTTCCAGGACAA 1901
DB 1139 CGTGTGCTGAATGGGAGCCGCGGTGATCACCGAGCGGAGGTGTCTGCTCCGTGAGGAGAGGGA 1138
QY 1902 GCACAGCTGTACTTTGTGATGGACTACATACACAGCGGGGATATGATGAGCCTGCTCAT 1961
DB 1199 GAATCTACCTGTACCTGGTCTGATGGAGTATTACGTGGCGGGGACCTGCTGACACTGCTGAG 1258
QY 1962 CAGGAT---GGAGGTCTTCCCTGAGCACCTGGCCGCTTCTACATTCGAGAGTTGACCT 2018
DB 1259 CAAGTTTGGGAGCGGATTCGCGCGGAGATGGCGGCTTCTACCTGGCGGAGATTGTCTAT 1318
QY 2019 GCCATTGAAAGTGTCCAAAGATGGCTTTATCCACGGGACATCAAGCCTTGACAACAT 2078
DB 1319 GGCCATAGACTCGGTGCACCGGCTTGGCTACGTGCACAGGGACATCAAACCCGACAACAT 1378
QY 2079 ACTCATCGACCTGGATGGTCAATTAAAGCTGACAGATTTTGGCCTCTGC 2127
DB 1379 CCTGCTGGACCGCTGTGGCCACATCCCGCTGGCGGACTTCGGCTCTTGC 1427

RESULT 13
US-09-938-842A-633
Sequence 633, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Krebs, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CON
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 633
LENGTH: 1398
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-633

Query Match 4.5%; Score 142; DB 9; Length 1398;
Best Local Similarity 57.4%;
Matches 256; Conservative Pred. No. 1.8e+28;
Mismatches 190; Indels 0; Gaps

QY 1690 TTGTGAAATCAAGACTCTAGGATCGGTCTTGGGAGTGCTGCTCTGTTGAAG 1749
DB 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111

; PRIOR FILING DATE: 2000-10-02
 ; NUMBER OF SEQ ID NOS: 3950
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2340
 ; LENGTH: 1635
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M34182
 US-09-880-107-2340

Query Match: 4.9%; Score 155.6; DB 10; Length 1635;
 Best Local Similarity 59.5%; Pred. No. 4e-32;
 Matches 263; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

Qy	1700	TCAGACTCTAGGCATCGCTTCCTTGGGGAAGTGTGCCCTCGCTTCTTAAGCTGGACACTC	1759
Db	192	TCAGGACGCTGGCATGGGCTCTCTTCGGGGGGTGTCTTGGTGGAGGACGACCG	251
Qy	1760	ACGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGG	1819
Db	252	CGGCCACTACGCCATGAAGATCTCTCAACAAGCAGAAGGTGGTGAAGATGAACAGGTGC	311
Qy	1820	CCATGTCAAGCTGAGAGGGACATCTCGCTGAAGCAGACAATGAGTGGTGGTCTCAAAAC	1879
Db	312	AGCACATACTGAACGAGAAGCGCATCTCTCGAGGCGATCGACTTCCGTTCTCGTCAAGC	371
Qy	1880	TCTACTACTCTTCCAGACAAGGACAGCCCTGTACTTTGGTATGGACTACATACCAGCGG	1939
Db	372	TCCAGTCTCTTTAAGGACAACTCCTACCTGTACTGTGATGGAGTACGTCCGGGTG	431
Qy	1940	GGGATATGATGAGCCTGTGTATCAGGATGGAGGTCTTCCCTGAGCACCTGGCCCGCTTCT	1999
Db	432	GGGAGATGTTTCTCCCGCCTACAGCGCGTCGGAAGGTTTAGCGAGCCCATGCGCTTCT	491
Qy	2000	ACATTGCAGAGTTGACCTCGGCGCATTGAAGTGTCCACAAGATGGGCTTTATCCACCGGG	2059
Db	492	ATGCCGCCAGGTCTGCTTGGCGTCCAGTACCTACACTCGCTGAGACCTATCCACCGCG	551
Qy	2060	ACATCAAGCCTGACACATACACTACGACCTGGATGGTGCATATTAAGCTGACAGATTTG	2119
Db	552	ACCTGAAGCCCGAGAATCTCTCATCGACGACGAGGCTACCTGCAGGTGACGACTTCG	611
Qy	2120	CGCTTGCCTGACTGGATTCAGGTG	2141
Db	612	GTTCGCCAAGCGCGTGAAGG	633

RESULT 12
 US-09-971-845-1
 ; Sequence 1, Application US/09971845
 ; Patent No. US20020132247A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DELANEY, ALLEN
 ; APPLICANT: YOGANATHAN, THILLAINATHAN
 ; TITLE OF INVENTION: DYSTROPHIA MYOTONICA PROTEIN KINASE
 ; TITLE OF INVENTION: (DM-PK) AND ITS USES
 ; FILE REFERENCE: KINE028CON
 ; CURRENT APPLICATION NUMBER: US/09/971,845
 ; CURRENT FILING DATE: 2001-10-04
 ; PRIOR APPLICATION NUMBER: PCT/US01/21479
 ; PRIOR FILING DATE: 2001-09-20
 ; PRIOR APPLICATION NUMBER: 60/238,558
 ; PRIOR FILING DATE: 2000-10-04
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3407
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-971-845-1

; Query Match 4.7%; Score 149.8; DB 10; Length 3407;

Db 400 TTTGAGGTTATGAAGGTTCTTGGGAAAGGTGCGTTTGGGAAAGCTTACCAGGTGAGGAA 459
 QY 1750 CTGGACACTCAGCCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCCTGAACCGG 1809
 Db 460 AAGGACACCTCTGAGATATACGCTATGAAGTATCAGGAGAGATCACATTTATGGAGAG 519
 QY 1810 AATCAAGTGGCCCATGTCAGGCTGAGAGGGACATCCTGGCTGAAGCAGACAAATGAGTGG 1869
 Db 520 AACCATGCTGAATACATGAAGCTGAGCGCGATATTTTACAAAAATTGATCACCCATTC 579
 QY 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACGCTCTACTTTTGTGATGACTAC 1929
 Db 580 ATGTTCAAACTTAAATACTCTTTTACAGACCAAGTATGATCTTGTGCTCGACTTT 639
 QY 1930 ATACGAGCGGGGATGATGAGCCTGCTGATCAGGATGAGGATGCTTCCCTGAGCACCTG 1989
 Db 640 ATAAAGGAGGTCATCTTTTCTTCCAACTCTATCACCAGGCTTTTCAGAGAGGACTTG 699
 QY 1990 GCCCGCTTCTATGTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAAGATGGGCTTT 2049
 Db 700 GCTGAGTGTACATGTCAGAAATCGTCTGCGAGTTTCCCATCTCCATCAGAAAGGCATA 759
 QY 2050 ATCCACCGGAGATCAAGCCTGACACATCTATCAGGATGAGGATGCTTCCCTGAGCACCTG 2109
 Db 760 ATGCACAGAGATCTGAAACCCGAAACATACATGATGATGAGGATGAGGATGAGGATG 819
 QY 2110 ACAGATTTTGGGCTCTGCACTGGATT 2135
 Db 820 ACAGATTTTGGTTAGCAAGAAAT 845

RESULT 14
 US-09-938-842A-2503
 ; Sequence 2503, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2503
 ; LENGTH: 1416
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-2503

Query Match 4.4%; Score 138.8; DB 9; Length 1416;
 Best Local Similarity 57.0%; Pred. No. 1.3e-27;
 Matches 254; Conservative 0; Mismatches 192; Indels 0; Gaps 0;
 QY 1690 TTTGTGAAATCAAGACTCTAGGATCGGTGCTTTGGGAAAGTGTGCTCCTGTTAG 1749
 Db 418 TTTGAGGTATTGAAGTTGTGGGACAAAGTGTGTTGAAAGTGTACCAGGTGAGGAA 477
 QY 1750 CTGGACACTCAGCCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCCTGAACCGG 1809
 Db 478 AAGACACGCTCTGAGATATACCGATGAAGTGTATGAGAAAGATATAAATTTGTTAGAG 537
 QY 1810 AATCAAGTGGCCCATGTCAGGCTGAGAGGGACATCCTGGCTGAAGCAGACAAATGAGTGG 1869
 Db 538 AATCATGCTGAATACATGAAGCCGAGCGGATATTTCTAACCAAAATCGATCATCTTTC 597

QY 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACGCTCTACTTTTGTGATGACTAC 1929
 Db 598 ATGTGCAACTTAAATACTCTTTTACAGACCAATACAGATTGTACTTCTGTTGACTTT 657
 QY 1930 ATACGAGCGGGGATGATGAGCCTGCTGATCAGGATGAGGATGCTTCCCTGAGCACCTG 1989
 Db 658 ATAAAGGAGGTCATCTTTTCTTCCAGCTCTATCACCAGGCTTTTTCAGGAGGACTTG 717
 QY 1990 GCCCGCTTCTATGTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAAGATGGCCTTT 2049
 Db 718 GCTGCTGTATACATGTCAGAAATCGTCTCTGCAAGTTTCCCATCTCCATGAGAAAGCATA 777
 QY 2050 ATCCACCGGAGATCAAGCCTGACACATCTATCAGGATGAGGATGCTTCCCTGAGCACCTG 2109
 Db 778 ATGCATAGAGATCTTAAACCTGAAACATCTATCAGGATGAGGATGAGGATGAGGATG 837
 QY 2110 ACAGATTTTGGGCTCTGCACTGGATT 2135
 Db 838 ACTGATTTTGGTTAGCAAGAAAT 863

RESULT 15
 US-09-764-868-58
 ; Sequence 58, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 58
 ; LENGTH: 1735
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-868-58

Query Match 4.3%; Score 135.6; DB 9; Length 1735;
 Best Local Similarity 58.3%; Pred. No. 1.1e-26;
 Matches 253; Conservative 2; Mismatches 176; Indels 3; Gaps 1;
 QY 1690 TTTGTGAAATCAAGACTCTAGGATCGGTGCTTTGGGAAAGTGTGCTCCTGTTAG 1749
 Db 297 TTTGAAATTAATTAAGTAATTTGGAAGAGTGTCTTTGGTGAAGTGTCTTCTCAAAATG 356
 QY 1750 CTGGACACTCAGCCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCCTGAACCGG 1809
 Db 357 AAGATCTGAAGCAATTTATGCAATGAAATCTCAACAACTGGGAGATGCTGAAGA 416
 QY 1810 AATCAAGTGGCCCATGTCAGGCTGAGAGGGACATCTCTGGCTGAAGCAGACAAATGAGTGG 1869
 Db 417 GCAGAGACCGCTGCTTCCGAGAGGAGCGCATGTGCTGTCAACGGGACTGCCAGTGG 476
 QY 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACGCTCTACTTTTGTGATGACTAC 1929
 Db 477 ATCACCAGGCTCTACATGAGCTTTCAGGAGGAGAACCCACTCTACTTACTGATGATGAT 536
 QY 1930 ATACGAGCGGGGATGATGAGCCTGCTGATCAGGATGAGGATGAGGATGAGGATGAGGAT 1986
 Db 537 TATGTGGTGTGATTTACTGACCTGCTCASCATAATTTGAAGACAAAGCTTCGGAAGAT 596
 QY 1987 CTGGCCCGCTTCTACATTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCCAAGATGGGC 2046
 Db 597 ATGCGGAGTCTACATTTGGTGAATTTGGTGTGCTGGCCATTGACTCCATCCATCAGCTTAT 656
 QY 2047 TTTATCCAGCGGACATCAAGCTTCAACATCTATCTGACCTGATGATGATGATGATGAT 2106
 Db 657 TACGKGACAGRGCATTAACCTGACATGTCCTTTTGGAGCGTGAATGATGATGATGATGAT 716

Qy 2107 CTGACAGATTTGG 2120
||| | || |||||
Db 717 CTGGCTGACTTTGG 730

Search completed: January 16, 2003, 22:04:38
Job time : 144.552 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 2727.25 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagccacccgaagtt.....aagagcattatttggggg 3155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	927.6	29.4	1547	11 BC024819	BC024819 Mus muscu
2	782.2	24.8	926	13 BI412714	BI412714 602987439
3	781	24.8	924	14 BQ930443	BQ930443 AGENCOURT
4	772	24.8	875	13 BI664597	BI664597 603290249
5	725	23.0	737	14 BM950668	BM950668 UI-M-EHOP
6	659.8	20.9	687	10 BB354465	BB354465 BB354465

7	637.4	20.2	930	14 BQ988648	BQ988648 AGENCOURT
8	622	19.7	689	13 BI331257	BI331257 602982912
9	578.8	18.3	824	13 BI525737	BI525737 602926563
10	562.4	17.8	575	12 BE986745	BE986745 UI-M-CG0P
11	553.8	17.6	569	12 BF461270	BF461270 UI-M-CG0P
12	540.6	17.1	549	10 AW494954	AW494954 UI-M-BH3-
13	537.4	17.0	539	12 BF465780	BF465780 UI-M-CG0P
14	533.8	16.9	537	10 AW456717	AW456717 UI-M-BH3-
15	519.4	16.5	863	10 BE620135	BE620135 601483053
16	508.6	16.1	515	13 BI133909	BI133909 UI-M-BH3-
17	503.6	16.0	590	13 BM052863	BM052863 1667C07.Y
18	503.6	16.0	957	14 BM801311	BM801311 AGENCOURT
19	497.6	15.8	631	9 AV253818	AV253818 AV253818
20	494.8	15.7	736	12 BG538495	BG538495 602567216
21	494.2	15.7	579	13 BI964175	BI964175 1665909.Y
22	479.4	15.2	596	12 BF137690	BF137690 601782626
23	476.4	15.1	478	12 BF460475	BF460475 UI-M-CG0P
24	471.4	14.9	473	12 BE987358	BE987358 UI-M-CG0P
25	458.8	14.5	462	12 BF463863	BF463863 UI-M-CG0P
26	443.6	14.1	462	12 BF465982	BF465982 UI-M-CG0P
27	442	14.0	450	9 AI430896	AI430896 mh70d12.x
28	438.6	13.9	445	10 AW493915	AW493915 UI-M-BH3-
29	429	13.6	686	10 BB614487	BB614487 BB614487
30	420.2	13.3	447	9 AA023402	AA023402 mh71d12.r
31	403.4	12.8	416	14 W81967	W81967 me93h03.r1
32	395.4	12.5	397	12 BF464894	BF464894 UI-M-CG0P
33	394.2	12.5	528	14 BM706472	BM706472 UI-E-CR1-
34	382.8	12.1	460	10 AW227865	AW227865 up15c06.Y
35	362	11.5	648	13 BI155504	BI155504 602903902
36	355.2	11.3	468	12 BF774005	BF774005 283544 MA
37	346.2	11.0	445	10 BB780786	BB780786 BB780786
38	343.8	10.9	355	10 AW909686	AW909686 ur72g09.Y
39	343	10.9	602	9 AL601079	AL601079 DKF2p313K
40	341.2	10.8	962	17 CNS0396F	AL233520 Tetraodon
41	338.2	10.7	343	12 BE853641	BE853641 ux20a08.Y
42	338	10.7	442	10 AW653741	AW653741 102578 MA
43	335	10.6	1053	14 U69566	U69566 U69566 Soar
44	333.4	10.6	488	10 BE289622	BE289622 601092179
45	332.6	10.5	419	10 AV655751	AV655751 AV655751

ALIGNMENTS

RESULT 1 BC024819 Mus musculus, Similar to large tumor suppressor 2, clone
LOCUS BC024819 1547 bp mRNA linear HTC 07-AUG-2002
DEFINITION IMAGE:5360539, mRNA.
ACCESSION BC024819
VERSION BC024819.2 GI:20380582
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1547)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
COMMENT On May 1, 2002 this sequence version replaced gi:19354132.
Contact: MGC help desk
Email: cgaps-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: ang@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 54 Row: m Column: 22
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7657294
 This clone has the following problem: no polyA-tail.

FEATURES

Location/Qualifiers
 1..1547
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5360539"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 350 a 468 c 446 g 283 t

BASE COUNT

ORIGIN

Query Match 29.4%; Score 927.6; DB 11; Length 1547;
 Best Local Similarity 99.6%; Pred. No. 1.8e-223;
 Matches 930; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGAGCCACCCGAGTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 60
 DB 614 ATGAGAGCCACCCGAGTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 673
 QY 61 CTCCTGCTTTTCCCAAGAGTCAGGCACTTCGGCAGCTGCAGAGGTGAACCGGCAGATG 120
 DB 674 CTCCTGCTTTTCCCAAGAGTCAGGCACTTCGGCAGCTGCAGAGGTGAACCGGCAGATG 733
 QY 121 CTTCAAGAGTTGGTAAGTCGCGCATGTGACAGAGATGCTGCAGAGGCTCACGAC 180
 DB 734 CTTCAAGAGTTGGTAAGTCGCGCATGTGACAGAGATGCTGCAGAGGCTCACGAC 793
 QY 181 ACGGCAGTAGAGTATGCAAGCTGCCTTGGAGTACATCAGTAGATGGGTACCTGAC 240
 DB 794 ACGGCAGTAGAGTATGCAAGCTGCCTTGGAGTACATCAGTAGATGGGTACCTGAC 853
 QY 241 CCCAGGATGACAGATGTCGAGTATCATCAAGCAGACCTCCCGAGAAAGGCGCTGGCG 300
 DB 854 CCCAGGATGACAGATGTCGAGTATCATCAAGCAGACCTCCCGAGAAAGGCGCTGGCG 913
 QY 301 TCCACCCCGGTGACTCGCGCGCCAGTTTCGAGGGCACAGGGAAGCACTCCCATCTAC 360
 DB 914 CCCACCCCGGTGACTCGCGCGCCAGTTTCGAGGGCACAGGGAAGCACTCCCATCTAC 973
 QY 361 CACAGCTGGGTGTGCAAACTACGAGGGCCCCCGCAGTGGAGAGATGCCGCGCA 420
 DB 974 CACAGCTGGGTGTGCAAACTACGAGGGCCCCCGCAGTGGAGAGATGCCGCGCA 1033
 QY 421 TATTAGACTTCTCTTCCCTGGAGCGGAGCGGACCCAGTGGAGATGCCGCGCA 480
 DB 1034 TATTAGACTTCTCTTCCCTGGAGCGGAGCGGACCCAGTGGAGATGCCGCGCA 1093
 QY 481 CATCTCCCAAGAGGTACAGCAGCAGATAGAGCAAGTGGCAGTTCGCGGCACACAC 540
 DB 1094 CATCTCCCAAGAGGTACAGCAGCAGATAGAGCAAGTGGCAGTTCGCGGCACACAC 1153
 QY 541 TATGGTCTGGTCTATCTATCTCGGAGAGTCGAGTATGGGTGTCAGCGCAGTTCCTCC 600
 DB 1154 TATGGTCTGGTCTATCTATCTCGGAGAGCTGGTATGGGTGTCAGCGCAGTTCCTCC 1213
 QY 601 TTCAGAAACAGAGCCAGCAGATGCTTATCCAGATGCGCAAGGCCAGGCTGGCCCT 660
 DB 1214 TTCAGAAACAGAGCCAGCAGATGCTTATCCAGATGCGCAAGGCCAGGCTGGCCCT 1273
 QY 661 CCGCGCAGCTACCTTTCTCTGCCCCATGCTGGGCTGTACACTGCCTCGCACCAAGCGC 720

DB 1274 CCGCCAGCCCTACCTTTCTGCCCCATCTGCTGGGTACACTGCTGCACCAAGCGC 1333
 QY 721 CGGGCTACCCACCTCGGGGCCACACCATTTACATGTGTGGGCACCCGGGTCACCGTTT 780
 DB 1334 CGGGCTACCCACCTCGGGGCCACACCATTTACATGTGTGGGCACCCGGGTCACCGTTT 1393
 QY 781 ACTGGCGAAAGCTCTGCACAGGCTGTGCTGCACCGCTCCAGGAACAGCCTCAATGCTGAC 840
 DB 1394 ACTGGCGAAAGCTCTGCACAGGCTGTGCTGCACCGCTCCAGGAACAGCCTCAATGCTGAC 1453
 QY 841 TTGTACGAGCTGGCTCCACAGTGCCTGCTGCTGCACCTCCACTGCGACCGCGGACTCG 900
 DB 1454 TTGTACGAGCTGGCTCCACAGTGCCTGCTGCTGCACCTCCACTGCGACCGCGGACTCG 1513
 QY 901 CTGCAAGAGCAGGCTCTAGAAGCCTCGCGCGCGC 934
 DB 1514 CTGCAAGAGCAGGCTCTAGAAGCCTCGCGCGCGC 1547

RESULT 2
 BI412714
 LOCUS 602987439F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5143496 5',
 DEFINITION mRNA sequence.
 ACCESSION BI412714
 VERSION BI412714.1 GI:15173637
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 926)
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone Sequencing by: Incyte Genomics, Inc.
 Distribution: NCI-CGAP: clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1353 row: n column: 09
 High quality sequence start: 4
 High quality sequence stop: 858
 Location/Qualifiers
 1..926
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5143496"
 /clone_lib="NCI_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a NotI - oligo(dT) primer [5'.
 TGTACCAATCTGAAGTGGAGGCGGCGCTCTGTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with NotI and cloned into the Not
 I and Eco RI sites of the modified pT73 vector. Library
 was constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source
 1..926
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5143496"
 /clone_lib="NCI_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a NotI - oligo(dT) primer [5'.
 TGTACCAATCTGAAGTGGAGGCGGCGCTCTGTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with NotI and cloned into the Not
 I and Eco RI sites of the modified pT73 vector. Library
 was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

ORIGIN

Query Match 24.8%; Score 782.2; DB 13; Length 926;
 Best Local Similarity 97.5%; Pred. No. 9e-187;

	Matches	890;	Conservative	0;	Mismatches	13;	Indels	10;	Gaps	9;
Qy	9	CACCCGAAAGTTTGGACCTTATCAAAAAGCTCTCAGGGAAATCGGATATCCCTCCTGCC	68							
Db	20	CACCCGAAAGTTTGGACCTTATCAAAAAGCTCTCA-GGAAATCCGATATCCCTCCTGCC	78							
Qy	69	TTTGTGCAACAGATCAGGCACCTTCGGCAGCTGCAGAGGTGAACCGGAGATGCTTTCAGGA	128							
Db	79	-TTTGTGCAACAGATCAGGCACCTTCGGCAGCTGCAGAGGTGAACCGGAGATGCTTTCAGGA	137							
Qy	129	GTTGGTGAATCGGCATGTGTACACAGGAGATGGCTGGCAGAGCGTTCACGACAGCGGGCAG	188							
Db	138	GTTGGTGAATCGGGATGTGACACAGGAGATGGCTGGCAGAGCGCTCAACGACAGCGGGCAG	197							
Qy	189	TAGGAGTATCGAAGCTGCCCTTGGAGTACATCAGTAAGATGGGCTACCTGGACCCAGGAA	248							
Db	198	TAGGAGTATCGAAGCTGCCCTTGGAGTACATCAGTAAGATGGGCTACCTGGACCCAGGAA	257							
Qy	249	TGACACAGATTGTGGAGTCATCAAGCAGACACCTCCACAGGAAGGGCCCTGGCGTTCACCCC	308							
Db	258	TGACACAGATTGTGGAGTCATCAAGCAGACACCTCCACAGGAAGGGCCCTGGCGCCACCCC	317							
Qy	309	GGTCACTCGGGCGGCCAGTTTTCAGAGGCGACAGGGGAAGCACTCCCATCTTACCACCACT	368							
Db	318	GGTGACTCGGGCGGCCAGTTTTCAGAGGCGACAGGGGAAGCACTCCCATCTTACCACCACT	377							
Qy	369	GGGTGGTGCAAACTACAGAGGCCCCCGCCACTCGAGAGAGATGCCGGCGCAATATTTAGA	428							
Db	378	GGGTGGTGCAAACTACAGAGGCCCCCGCCACTCGAGAGAGATGCCGGCGCAATATTTAGA	437							
Qy	429	CTTTCTCTTCCCTGGAGCGCGGACCCACCGTGGCCAGGCTCACAGCATCTCTCC	488							
Db	438	CTTTCTCTTCCCTGGAGCGCGGACCCACCGTGGCCAGGCTCACAGCATCTCTCC	497							
Qy	489	CAAAGGTTACAGCACAGCATAGAGCAAGTGCGCACTTTCCGGGCACACACTATGGTCG	548							
Db	498	CAAAGGTTACAGCACAGCATAGAGCAAGTGCGCACTTTCCGGGCACACACTATGGTCG	557							
Qy	549	TGCTCATCTACTATCGAGCAGCTTGGGTATGGGGTGCAGCGCATGCTCTCTCTCCAGAA	608							
Db	558	TGCTCATCTACTATCGAGCAGCCTGGGTATGGGGTGCAGCGCATGCTCTCTCTCCAGAA	617							
Qy	609	CAAGACGCCACCATGCCTATTCAGCATGGCCAAAGGCCAGGGTGGCCCTCCGCGCAG	668							
Db	618	CAAGACGCCACCATGCCTATTCAGCATGGCCAAAGGCCAGGGTGGCCCTCCGCGCAG	677							
Qy	669	CCTCACCTTTCTGCCCATGCTGGGCTGTACACTGCCTCGCACCC-ACAAGCCGGCGGCTA	727							
Db	678	CCTCACCTTTCTGCCCATGCTGGGCTGTACACTGCCTCGCACCCAAAGCCGGCGGTTA	737							
Qy	728	CCCCACCTGGGGCCCCCAATACATGTGTTGGCACCCCGGGGTCCCACGTTTACTGSGC	787							
Db	738	CCCCACCTGGGGCCCCCAATACATGTGTTGGCACCCCGGGGTCCCACGTTTACTGSGC	795							
Qy	788	AAAGCTCTGCACAGGCTGTGCTGCGACCGTCCAGGAACAGCCTCAATGCTCACTTGACG	847							
Db	796	AAAGCTCTGCACAGGCTGTGCTGCGA-CGTCACAGGAACAG-CTCAATGCTCACTTGACG	853							
Qy	848	-AGCTGGGCTCCAC-GGTGCCCTGGTC-TGCAGCTCCACTGGCGACCGCGCACTTCGGTGC	904							
Db	854	AGCTGGGCTCCACGGGTGCCCTGGTCTTGAGGTCCCACTGGCACCGCCGGAATCGCTGC	913							
Qy	905	AGAGCAGGGTCT	917							
Db	914	AGAGCAGGGTCT	926							

RESULT 3
BQ930443
LOCUS
DEFINITION: 924 bp mRNA linear EST 20-AUG-2002
AGENCOURT:8931059 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6468340
5', mRNA sequence.
ACCESSION BQ930443

BQ930443.1 GI:22345474
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 924)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-re@mail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 DNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L7AM13995 row: 1 column: 05
 High quality sequence stop: 554.
 Location/Qualifiers
 1. 924
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:6468340"
 /clone_lib="NIH_MGC_94"
 /tissue_type="retina"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 3.3 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."
 230 a 257 c 241 g 196 t
 BASE COUNT
 ORIGIN

	Query Match	24.8%	Score 781;	DB 14;	Length 924;
	Best Local Similarity	99.3%	Pred. No. 1.8e-186;		
	Matches 795;	Conservative 0;	Mismatches 5;	Indels 1;	Gaps 1;
QY	2354	GGTACAGCAGCTCTGTGACTGCTGGAGCGCTGGTGATCTCTTTGAGATGCTGGTTG	2413		
Db	7	GGGACACGACGCTCTGTGACTGCTGGAGCGCTGGTGATCTCTTTGAGATGCTGGTTG	66		
QY	2414	GGCAGCGCCCTTTCTTGGCCGCCCCACCCCCACAGAGACGAGCTGAAGGTGATCAACTGGG	2473		
Db	67	GGCAGCGCCCTTTCTTGGCCGCCCCACCCCCACAGAGACGAGCTGAAGGTGATCAACTGGG	126		
QY	2474	AGAGCAGCTGCATATCCCTACGAGGTGAGGCTCAGGCTGAGCCCGAGACCTCATCA	2533		
Db	127	AGAGCAGCTGCATATCCCTACGAGGTGAGGCTCAGGCTGAGCCCGAGACCTCATCA	186		
QY	2534	CGAAGCTGTGCTGGCGGGCTGACTGCCGCTGGCAGGATGGGCGAGATCACCTCAAGG	2593		
Db	187	CGAAGCTGTGCTGGCGGGCTGACTGCCGCTGGCAGGATGGGCGAGATCACCTCAAGG	246		
QY	2594	CACACCCGTTCTTCAACACCATCGACTTTTCCCGTGACATCCGAAAGCAGGCTGCACCT	2653		
Db	247	CACACCCGTTCTTCAACACCATCGACTTTTCCCGTGACATCCGAAAGCAGGCTGCACCT	306		
QY	2654	ACGTCCCCACCATCAGCCACCCCATGGACACCTCCAAATTTTGACCCGCTGATGAAGAA	2713		
Db	307	ACGTCCCCACCATCAGCCACCCCATGGACACCTCCAAATTTTGACCCGCTGATGAAGAA	366		
QY	2714	GCCTCTGGCAGCGGCCACGCGAGAGCGCAAGGCTGGACACGCTGGCCTCCGCCA	2773		
Db	367	GCCTCTGGCAGCGGCCACGCGAGAGCGCCACGAGGCTGGGACACGCTGGCCTCCGCCA	426		
QY	2774	GCACAAGCATCCAGAGCAGCCCTTCTATGAGTTTCACCTTCCCGCAGGTTCTTCGATGACA	2833		
Db	427	GCACAAGCATCCAGAGCAGCCCTTCTATGAGTTTCACCTTCCCGCAGGTTCTTCGATGACA	486		

QY 2834 ACGGCTATCCCTCCGGTGCCGGAAGCCCTCAGAGCCCGCAGAGAGTGACAGCCCGAGGG 2893
 Db 487 ACGGCTATCCCTCCGGTGCCGGAAGCCCTCAGAGCCCGCAGAGAGTGACAGCCCGAGGG 546
 QY 2894 ATGGGGACTTGGAGGTGGCGGCGAGGGCTGCCAGCGGTGTACGTGTAGACCTCAGTTA 2953
 Db 547 ATGGGGACTTGGAGGTGGCGGCGAGGGCTGCCAGCGGTGTACGTGTAGACCTCAGTTA 606
 QY 2954 ACCAACAACCTGAGGAAACCCAAATAGATTTCTTTTCAAGACAACTCAAGCTTAGG 3013
 Db 607 ACCAACAACCTGAGGAAACCCAAATAGATTTCTTTTCAAGACAACTCAAGCTTAGG 666
 QY 3014 AATCCTTCAATTTTGTCTGTAATGCGCAACAGAGAGTCAACATGATTTCAAATT 3073
 Db 667 AATCCTTCAATTTTGTCTGTAATGCGCAACAGAGTCAACATGATTTCAAATT 726
 QY 3074 AGCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAAATTAGTACAGTATG 3133
 Db 727 AGCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAAATTAGTACAGTATG 785
 QY 3134 GAAAGACCACTTATTTTGGG 3154
 Db 786 GAAAGACCACTTATTTTGTG 806

RESULT 4
 BI664597 875 bp mRNA linear EST 12-SEP-2001
 LOCUS 603290249F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5324077 5',
 DEFINITION mRNA sequence.
 ACCESSION BI664597
 VERSION BI664597.1 GI:15578830
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NTH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cyabps-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM1822 row: b column: 14
 High quality sequence stop: 798.

FEATURES
 source
 1..875
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5324077"
 /clone.lib="NCI_CGAP_Mam6"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies, Investigator
 providing samples: Jeffrey Green, M.D., NIH"
 205 a 267 c 248 g 155 t
 BASE COUNT 205 a 267 c 248 g 155 t
 ORIGIN
 Query Match 24.5%; Score 772; DB 13; Length 875;
 Best Local Similarity 99.08; Pred. No. 3.3e-184;
 Matches 798; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 2128 ACTGGATTGAGTGGAGTCCAAATTCACAAATTCACAAAGAGGAAACACATGAGACAG 2187
 Db 1 ACTGGATTGAGTGGAGTCCAAATTCACAAATTCACAAAGAGGAAACACATGAGACAG 60
 QY 2188 GACAGATGGAGCCCGGTGACCTCTGGGAGCATGTTTCCAACTGCTGCTGGAGACAGG 2247
 Db 61 GACAGATGGAGCCCGGTGACCTCTGGGAGCATGTTTCCAACTGCTGCTGGAGACAGG 120
 QY 2248 TTAAGACCTGAGCAGAGGCGCAGAGCAGCAGAGTGCCTGGCACAATTCCTTT 2307
 Db 121 TTAAGACCTGAGCAGAGGCGCAGAGCAGCAGAGTGCCTGGCACAATTCCTTT 180
 QY 2308 GTCGGGACACCAAAATTCATGCTGCTCCGAGGTGCTTCTCCGAAAGGGTACAGCAGCTC 2367
 Db 181 GTCGGGACACCAAAATTCATGCTGCTCCGAGGTGCTTCTCCGAAAGGGTACAGCAGCTC 240
 QY 2368 TGTGACTGTTGGAGCGTGGTGTGATTCTTTGAGATGCTGCTGGGAGCGCTTTC 2427
 Db 241 TGTGACTGTTGGAGCGTGGTGTGATTCTTTGAGATGCTGCTGGGAGCGCTTTC 300
 QY 2428 TTGGCCCCACCCACACAGAGCAGCAGTGAAGTGAATCACTGGGAGAGCAGCTGCAT 2487
 Db 301 TTGGCCCCACCCACACAGAGCAGCAGTGAAGTGAATCACTGGGAGAGCAGCTGCAT 360
 QY 2488 ATCCCTACGAGTGGAGTTCAGCGTTCAGGCGGCGAGACCTCATCAGAGCTGTGCTGC 2547
 Db 361 ATCCCTACGAGTGGAGTTCAGCGTTCAGGCGGCGAGACCTCATCAGAGCTGTGCTGC 420
 QY 2548 GCGGCTGACTGCGGCTGGGCGAGGATGGGCGAGTCAAGCAGCAGCCGCTTCTTC 2607
 Db 421 GCGGCTGACTGCGGCTGGGCGAGGATGGGCGAGTCAAGCAGCAGCCGCTTCTTC 480
 QY 2608 AACACCATGACTTTTCCCGTGACATCCGAAAGCAGGCTGCACCTTACGTCCTCCACATC 2667
 Db 481 AACACCATGACTTTTCCCGTGACATCCGAAAGCAGGCTGCACCTTACGTCCTCCACATC 540
 QY 2668 AGCCACCCCTGAGCAGCTCCAAATTTTACCCGCTGATGAAGAAAGCCCTGGCAGGAG 2727
 Db 541 AGCCACCCCTGAGCAGCTCCAAATTTTACCCGCTGATGAAGAAAGCCCTGGCAGGAG 600
 QY 2728 GCCAGCGGAGAGCGCCCAAGGCTGGGACACGCTGGGCTCCCGCTCCCGCAGCAGCATCCA 2787
 Db 601 GCCAGCGGAGAGCGCCCAAGGCTGGGACACGCTGGGCTCCCGCAGCAGCATCCA 660
 QY 2788 GAGCAGCCCTTCTATGATTCACCTTCCGAGGTTCCTTCGATGACACAGGCTATCCCTTC 2847
 Db 661 GAGCAGCCCTTCTATGATTCACCTTCCGAGGTTCCTTCGATGACACAGGCTATCCCTTC 720
 QY 2848 CGGTGCGGAGCCCTCAGAGCCCGCAGAGTGCAGACCCAGGGG--ATGCGGACTTGG 2905
 Db 721 CGGTGCGGAGCCCTCAGAGCCCGCAGAGTGCAGACCCAGGGGATGCGGACTTGG 780
 QY 2906 AAGGTGCGGCGGAGGCTGCCACCG 2931
 Db 781 AAGGTGCGG--CGAGGCTGCCACCG 805

RESULT 5
 BM950668 737 bp mRNA linear EST 14-MAR-2002
 LOCUS UT-M-EP0p-buu-b-12-0-UI.r1 NIH_BMAP_EH0p Mus musculus cDNA clone
 DEFINITION IMAGE:5686955 5', mRNA sequence.
 ACCESSION BM950668
 VERSION BM950668
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NTH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers
1. 737
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5686955"
/clone_lib="NIH_BMAP_EH0p"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: Brain; Vector: pYX-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according to
Bonald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). "Gene Discovery in the Developing Mouse Nervous
System", supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT
ORIGIN

Query Match 23.0%; Score 725; DB 14; Length 737;

Best Local Similarity 99.7%; Pred. No. 2.4e-172;

Matches 736; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1370 CGAAGGGAGACAAAGCTGCGAGAGACAAAGACGATTCAGACCTCCCGGTGCGCTGTC 1429
DB 1 CGAAGGGAGACAAAGCTGCGAGAGACAAAGACGATTCAGACCTCCCGGTGCGCTGTC 60
QY 1430 CCAAGAATAGCAGATGAAGAAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATG 1489
DB 61 CCAAGAATAGCAGATGAAGAAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATG 120
QY 1490 CTTCAATTTCTATGAGCAACAGCTGGAGAAATGTCATCAAACTTACCAGCAGAGG 1549
DB 121 CTTCAATTTCTATGAGC-ACACGTGGAGAAATGTCATCAAACTTACCAGCAGAGG 179
QY 1550 TCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGCTGTGAGGCCGAGC 1609
DB 180 TCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGCTGTGAGGCCGAGC 239
QY 1610 AGGACGATGAGGAAGATCTCTTACCAGAGGAGTCTAACTTACAAACCGCTTGAAGAGG 1669
DB 240 AGGACGATGAGGAAGATCTCTTACCAGAGGAGTCTAACTTACAAACCGCTTGAAGAGG 299
QY 1670 CCAAGATGGAAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 1729
DB 300 CCAAGATGGAAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 359
QY 1730 AAGTGTGCTGCTTGTAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGA 1789
DB 360 AAGTGTGCTGCTTGTAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGA 419

QY 1790 AGAAGGATCTCTGAAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGACATCTCG 1849
DB 420 AGAAGGATCTCTGAAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGACATCTCG 479
QY 1850 CTGAAGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGGACAAGCAGCC 1909
DB 480 CTGAAGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGGACAAGCAGCC 539
QY 1910 TGTACTTTGTGATGGACTACATACCAAGCGGGGATATGATGAGCCCTGCTGATCAGATGG 1969
DB 540 TGTACTTTGTGATGGACTACATACCAAGCGGGGATATGATGAGCCCTGCTGATCAGATGG 599
QY 1970 AGTCTTCTCCCTGAGCACCTGGCCCGCTTCTACATTCAGAGTTGACCTGGCCATTGAA 2029
DB 600 AGTCTTCTCCCTGAGCACCTGGCCCGCTTCTACATTCAGAGTTGACCTGGCCATTGAA 659
QY 2030 GTGTCCACAAGATGGGCTTTTATCCACCGGAGATCAAGCTGACAACATCTACTCAGCC 2089
DB 660 GTGTCCACAAGATGGGCTTTTATCCACCGGAGATCAAGCTGACAACATCTACTCAGCC 719
QY 2090 TGGATGGTCATATTAGC 2107
DB 720 TGGATGGTCATATTAGC 737

RESULT 6

BB354465

LOCUS

BB354465

DEFINITION

BB354465

ACCESSION

BB354465

VERSION

BB354465.2

KEYWORDS

EST.

SOURCE

Mus musculus

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 687)

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

TITLE

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL

Unpublished (2001)

COMMENT

On Jul 12, 2000 this sequence version replaced gi:9066293.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp.

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayashizaki,Y., Sugahara,Y., Itoh

M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura

S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

BB354465 687 bp mRNA linear EST 24-OCT-2001
BB354465 RIKEN full-length enriched, 10 days neonate cerebellum Mus
musculus cDNA clone B930098K05 3', mRNA sequence.

ACCESSION BB354465

VERSION BB354465.2

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 687)

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL Unpublished (2001)

COMMENT On Jul 12, 2000 this sequence version replaced gi:9066293.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp.

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayashizaki,Y., Sugahara,Y., Itoh

M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura

S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a


```

QY 1919 TGATGGACTACATACAGCGGGGATATGATGAGCCTGCTGCTATCAGGATGAGAGTCTTCC 1978
Db 551 TGATGGACTACATACAGCGGGGATATGATGAGCCTGCTGCTATCAGGATGAGAGTCTTCC 610
QY 1979 CTGAGCACCTGGCCGCTTCTACATTCAGAGTTGACCTGGCCATTGAAAGTGTCCACA 2038
Db 611 CTGAGCACCTGGCCGCTTCTACATTCAGAGTTGACCTGGCCATTGAAAGTGTCCACA 670
QY 2039 AGATGGGCTTTATCCA 2054
Db 671 AGATGGGCTTTATCCA 686

RESULT 9
BI525737/c
LOCUS 602926563F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5059023 5',
DEFINITION mRNA sequence.
ACCESSION BI525737
VERSION BI525737.1 GI:15350529
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL 1 (bases 1 to 824)
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A11159 row: n column: 16
High quality sequence start: 10
High quality sequence stop: 745.

FEATURES
source Location/Qualifiers
1..824
/organism="Mus musculus"
/strain="Czech II"
/db_xref="taxon:10090"
/clone="IMAGE:5059023"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: p7T3D-Pac (Pharmacia) with a
modified polylinker; Site.1: NotI; Site.2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'].
TGTTACCAATCTGAAGTGGGCGGCTCTGTTTGTCTTTTCTTTT 3' ].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified p7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 178 a 232 c 219 g 195 t
ORIGIN

Query Match 18.3%; Score 578.8; DB 13; Length 824;
Best Local Similarity 94.3%; Pred. No. 2.5e-135;
Matches 711; Conservative 0; Mismatches 27; Indels 16; Gaps 10;

QY 1722 CTTTGGGGAAGTGCCTCGCTGCTTAAGCTGGACA-----CTCAGCGTCTGTAGCCCATG 1776
Db 824 CTTTGGGGAAGTGCCTCGCTGCTTGTAACTGGGACACTGCACGCTTCTGTAGCCCATG 765
QY 1777 AAG--ACTCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGGCCCAT--GTCAAGGC 1832

```

```

Db 764 AAGGACTCTCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGGCCCAATGGTCAAGCC 705
QY 1833 TGAGAGGACATCTCTGGCTGAAGCAGACAAT-GAGTGGGTGGTCAAACTC-TACTACTCC 1890
Db 704 TGAGAGGACATCTCTGGCTGAAGCAGACCATTGGAGTGGGTGCTCAAACCTCTTACTACTCC 645
QY 1891 T-TCCAGSACAAGACAGACGCTGTACTTTTGTGATGAGCTATACATACAGCGGGGATATGAT 1949
Db 644 TGTCCAGSACAAGACAGACGCTGTACTTTTGTGATGAGCTATACATA-CAGCGGGGATATGAT 586
QY 1950 GAGCCTGCTGATCAGGATGAGAGTCTTCCCTGAGAGACCTGCGCCGCTTCTACATTTGACAG 2009
Db 585 GAGCCTGCTGATCAGGATGAGAGTCTTCCCTGAGAGACCTGCGCCGCTTCTACATTTGACAG 526
QY 2010 -GTTGACCTTGGCCATTGAAAGTGTCCACAAGATGGGCTTTATCCACGGGACATCAAGC 2068
Db 525 GGTTGACCTTGGCCATTGAAAGTGTCCACAAGATGGGCTTTATCCACGGGACATCAAGC 466
QY 2069 CTGACAACATACTCATCGACCTGGATGCTCATATTAAGCTGACAGATTTTGGCCTCTGCA 2128
Db 465 CTGACAACATACTCATCGACCTGGATGCTCATATTAAGCTGACAGATTTTGGCCTCTGCA 406
QY 2129 CTGATTCAGTGGATCTACAAATTCCTCAAGTACTACAGAAAGGACACATGACAGAG 2188
Db 405 CTGATTCAGTGGATCTACAAATTCCTCAAGTACTACAGAAAGGACACATGACAGAG 346
QY 2189 ACACATGGAGCCGCTGACCTCTGGGACGATGTTTCCAACTGTCGCTGTGGAGACAGGT 2248
Db 345 ACACATGGAGCCGCTGACCTCTGGGACGATGTTTCCAACTGTCGCTGTGGAGACAGGT 286
QY 2249 TAAAGACCTTGGAGCAGAGCGGCGAGAGCAGACAGAGGTGCTGSCACATTTCTCTTG 2308
Db 285 TAAAGA -CTTGGAGCAGAGCGGCGAGAGCAGACAGAGGTGCTGSCACATTTCTTGT 227
QY 2309 TCGGACACCAAAATACATCTCGGAGGTGCTTCTCCGAAAGGTTACAGCAGCTCT 2368
Db 226 TCGGACACCAAAATACATCTCGGAGGTGCTTCTCCGAAAGGTTACAGCAGCTCT 167
QY 2369 GTGACTGTGTGAGCGTGTGCTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTCT 2428
Db 166 GTGACTGTGTGAGCGTGTGCTGATTTCTTTGAGATGCTGTTGGGAGCGGCTTTCT 108
QY 2429 TGGCCCCACCCACAGACAGACGAGTGAAGT 2462
Db 107 TGGCCCCACCCACAGACAGACGAGTGAAGT 74

RESULT 10
BE986745/c
LOCUS BE986745 575 bp mRNA. linear EST 29-APR-2002
DEFINITION UI-M-CGOp-bey-c-06-0-UI.s1 NIH-BMAP_Ret4_S2 Mus musculus cDNA clone
UI-M-CGOp-bey-c-06-0-UI 3', mRNA sequence.
ACCESSION BE986745
VERSION BE986745.1 GI:10661386
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 575)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov

```


Oligo-3T track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
Seq primer: M13 Forward
POLYA-No.

FEATURES
source

Location/Qualifiers
1..575
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bey-c-06-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_Ret4_S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
TAG_SEQ=None found"

BASE COUNT 102 a 163 c 200 g 110 t

Query Match 17.8%; Score 562.4; DB 12; Length 575;
Best Local Similarity 99.7%; Pred. No. 3e-131;
Matches 574; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 358 TACCACAGCTGGGTGCTCAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGG 417
DB 575 TACCACAGCTGGGTGCTCAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGG 516
QY 418 CAATATTTAGACTTCTTCTCCCTGGAGCGCGGAGCCAGCCAGCGTCCAGGCTAC 477
DB 515 CAATATTTAGACTTCTTCTCCCTGGAGCGCGGAGCCAGCCAGCGTCCAGGCTAC 457
QY 478 CAGCATCTCTCCAAAGGGTACAGCACAGCAGTAGAGCAAGTGGCACTTTCCGGGCACA 537
DB 456 CAGCATCTCTCCAAAGGGTACAGCACAGCAGTAGAGCAAGTGGCACTTTCCGGGCACA 397
QY 538 CACTATGGTGGTGTCACTACTATCGGAGCAGTCTGGGTATGGGTGCGAGCGCATTC 597
DB 396 CACTATGGTGGTGTCACTACTATCGGAGCAGCCTGGGTATGGGTGCGAGCGCATTC 337
QY 598 TCCTTCCAGAACAGCGCCACCATGCTATTCAGCATGCCAAGCCAGCGGTGGC 657
DB 336 TCCTTCCAGAACAGCGCCACCATGCTATTCAGCATGCCAAGCCAGCGGTGGC 277
QY 658 CTTCCCGGCGAGCTTCACCTTTCTCCCATGCTGGGTGTACACTGCTCGCACCAAG 717
DB 276 CTTCCCGGCGAGCTTCACCTTTCTCCCATGCTGGGTGTACACTGCTCGCACCAAG 217
QY 718 CCGCGGCTACCCACCTGGGCGCCACCATATACATGTTGGGACCGCGGGTCCCGAG 777
DB 216 CCGCGGCTACCCACCTGGGCGCCACCATATACATGTTGGGACCGCGGGTCCCGAG 157
QY 778 TTTACTGGGAAGCTCTGCACAGCTGTGTGGCAGCCGTCAGGAAGAGCTCAATGCT 837
DB 156 TTTACTGGGAAGCTCTGCACAGCTGTGTGGCAGCCGTCAGGAAGAGCTCAATGCT 97
QY 838 GACTTGTAGAGCTGGGTCCACAGTGTGCTGCTGAGCTCCAGCTGCGAGCGCGCAG 897
DB 96 GACTTGTAGAGCTGGGTCCACAGTGTGCTGCTGAGCTCCAGCTGCGAGCGCGCAG 37
QY 898 TCGTGCAGAGCAGGGTCTAGAGAGCTCGCGCGG 933

Db 36 TCGCTGCAGAGCAGGTCTAGAGCCTCGCGCGG 1

RESULT 11
BF461270/c

LOCUS
DEFINITION

ACCESSION
BF461270

VERSION
BF461270.1

KEYWORDS
EST.

ORGANISM
Mus musculus

REFERENCE
1 (Bases 1 to 569)

AUTHORS:
Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE
Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL
Genome Res. 6 (9), 791-806 (1996)

MEDLINE
97044477

COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20992-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov

FEATURES
Source

Location/Qualifiers
1..569
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bms-a-06-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_Ret4_S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
TAG_SEQ=None found"

BASE COUNT 101 a 161 c 200 g 107 t

Query Match 17.6%; Score 553.8; DB 12; Length 569;
Best Local Similarity 99.5%; Pred. No. 4.6e-129;
Matches 566; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 367 CTGGGTGGTGCACAACTA-CGAGGGCCCGCGCGAGAGATGCGCGGCAATATTT 425
DB 569 CTGGGTGGTGCACAACTA-CGAGGGCCCGCGCGAGAGATGCGCGGCAATATTT 510
QY 426 AGACTTTCTTCTTCCCTGGAGCGGAGCCAGCCAGCGTCCAGGCTCACCAGCATCC 485
DB 509 AGACTTTCTTCTTCCCTGGAGCGGAGCCAGCCAGCGTCCAGGCTCACCAGCATCC 450
QY 486 TCCCAAGGGTACAGCAGCAGTAGAGCAGTGGCAGCTTTCCGGGGCACACATATGG 545


```

QY 1468 ATCAAGAGT 1476
Db 541 ATCAAGAGT 549

RESULT 13
BF465780/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA-No.
Location/Qualifiers
1..539
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bqr-g-12-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu. The tissue
for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_SEQ=None found"
97 a 151 c 189 g 102 t

BASE COUNT.
ORIGIN

Query Match
Best Local Similarity 17.0%; Score 537.4; DB 12; Length 539;
Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 396 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCCCTGGAGCCGAGCCGG 455
Db 539 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCTCCCTGGAGCCGAGCCGG 480
QY 456 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 515
Db 479 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 420
QY 516 AAGTGGCGCACTTTCGGGGCACACACTATGCTGCTGCTACTATCGGAGCAGTCTGG 575

QY 419 AAGTGGCGCACTTTCGGGGCACACACTATGCTGCTGCTACTATCGGAGCAGCTGG 360
QY 576 GTATGGGGTGCAGCGCAGTTCCTTCCAGAACAGAGCCACAGATGCTATTCCAG 635
Db 359 GTATGGGGTGCAGCGCAGTTCCTTCCAGAACAGAGCCACAGATGCTATTCCAG 300
QY 636 CATGGCAAGGCCAGGGTGGCCCTCCCGCCAGCCCTACCTTTCCTGGCCATGCTGGGCT 695
Db 299 CATGGCAAGGCCAGGGTGGCCCTCCCGCCAGCCCTACCTTTCCTGGCCATGCTGGGCT 240
QY 696 GTACACTGCTCGCACCAAGCCGCGGCTACCCACCTGGGGCCACCCATACATGT 755
Db 239 GTACACTGCTCGCACCAAGCCGCGGCTACCCACCTGGGGCCACCCATACATGT 180
QY 756 GTTGGGACCCCGGGTCCACAGCTTTACTGGGAAAGCTTGCACAGGCTGTGCTGGCACC 815
Db 179 GTTGGGACCCCGGGTCCACAGCTTTACTGGGAAAGCTTGCACAGGCTGTGCTGGCACC 120
QY 816 GTCCAGGAACAGCCTCAATGCTGACTTGTACAGCTGGGCTCCAGGTCCTGGTCTGC 875
Db 119 GTCCAGGAACAGCCTCAATGCTGACTTGTACAGCTGGGCTCCAGGTCCTGGTCTGC 60
QY 876 AGCTCCACTGGCAGCCGCGACTCGCTGCAGAGCAGGCTCTAGAGCCTCGCGGCCG 934
Db 59 AGCTCCACTGGCAGCCGCGACTCGCTGCAGAGCAGGCTCTAGAGCCTCGCGGCCG 1

RESULT 14
AM456717/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: MEST@mail.nih.gov
Oligo-dT track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA-No.
Location/Qualifiers
1..539
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bqr-g-12-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
NIH_BMAP_Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu. The tissue
for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_SEQ=None found"
97 a 151 c 189 g 102 t

BASE COUNT.
ORIGIN

Query Match
Best Local Similarity 17.0%; Score 537.4; DB 12; Length 539;
Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 396 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCCCTGGAGCCGAGCCGG 455
Db 539 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCTCCCTGGAGCCGAGCCGG 480
QY 456 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 515
Db 479 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 420
QY 516 AAGTGGCGCACTTTCGGGGCACACACTATGCTGCTGCTACTATCGGAGCAGTCTGG 575

```

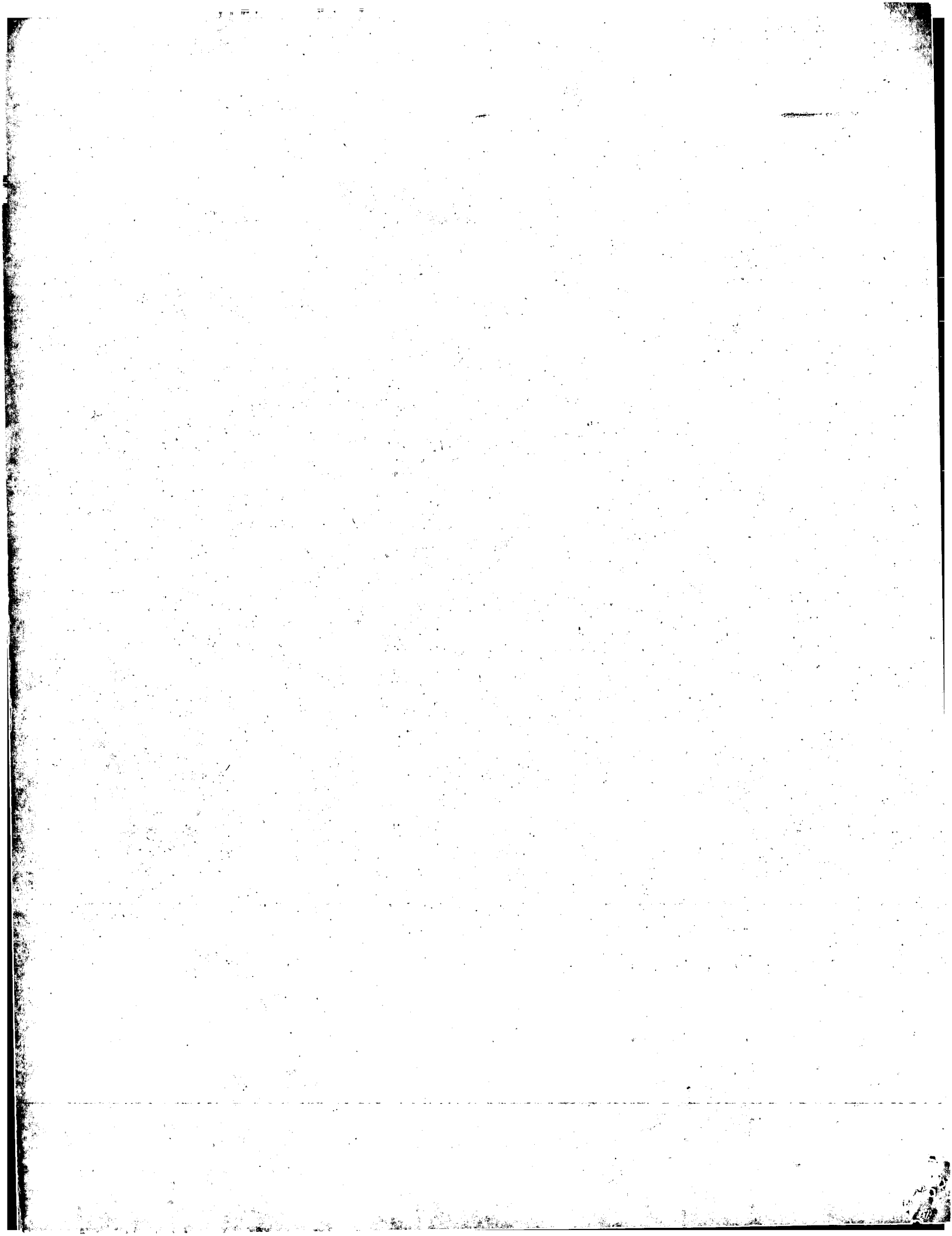


```

||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 426 CAGATTTCGGCCTCTGCACTGGGTTTCAGGTGGACTCACAATTCAAAATATTACCAGAAAG 485
QY 2171 GGAACACATGACAGACAGCATGGAGCCGGTGACCTCTGGGACGATGTTTCCAAC 2230
Db 486 GGAGCCATGTCAGACAGGACAGCATGGAGCCCGACGACCTCTGGGATGATGTCTAACT 545
QY 2231 GTCGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCGCAGAAAGCAGCAGAGGT 2290
Db 546 GTCGGTGTGGGACAGGCTGAAGACCCTAGACAGAGGGCGCGAAGCAGCAGAGGT 605
QY 2291 G--CCTGGCACATTCCTTGTGGGACACAAA-TTACATCGCTCCGGAGGTCTTC--T 2345
Db 606 GGCCGTGGCACATTCACCTGTGGGACTCCAAACTAACATCGCACCCGAGGTCTCCCTC 665
QY 2346 CCGCAAGGGTACACGACGCTCTGTGACTGGTG 2378
Db 666 CGCAAGGGGTACACTCAACTCTGTGACTGGGG 698

```

Search completed: January 16, 2003, 21:49:12
Job time : 2750.25 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 9497.51 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1 atctagcagcgagcgagcaaa.....gcttggccaattgatttac 5720

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sv.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5708	99.8	5720	3	DMU29608	U29608 Drosophila
2	5708	99.8	5720	6	AR201454	AR201454 Sequence
3	5216.4	91.2	5360	6	AR201454	L39837 Drosophila
c	3969	69.4	163741	3	AE007821	AE007821 Drosophila
c	3942.2	68.9	236109	3	AE003775	AE003775 Drosophila
c	2965.2	51.8	28133	2	AC012975	AC012975 Drosophila
c	981.2	17.2	39839	2	AC013000	AC013000 Drosophila
8	567.8	9.9	3155	6	AR201457	AR201457 Sequence
9	565.4	9.9	3460	10	AB023958	AB023958 Mus muscu
10	558.6	9.8	5486	6	E38226	E38226 Human tumor
11	558.6	9.8	5486	6	E38227	E38227 Human tumor
12	557	9.7	3412	9	AB028019	AB028019 Homo sapi
13	557	9.7	4098	9	AF207547	AF207547 Homo sapi
14	522.8	9.1	3213	6	AR201456	AR201456 Sequence
15	522.8	9.1	3213	10	AF104414	AF104414 Mus muscu
16	470.6	8.2	3984	6	AR201455	AR201455 Sequence
17	470.6	8.2	4241	9	AF164041	AF164041 Homo sapi
18	470.6	8.2	4307	9	AF104413	AF104413 Homo sapi
19	469	8.2	2442	6	E24613	E24613 warts prote
20	455.8	8.0	1374	6	E24614	E24614 warts prote
c	307.6	5.4	181750	2	AC098623	AC098623 Rattus no
c	303.2	5.3	183597	9	AL356285	AL356285 Human DNA
23	301.2	5.3	164490	2	AC102712	AC102712 Mus muscu
24	297.4	5.2	2231	8	AF041843	AF041843 Ustilago
25	290.8	5.1	2126	3	AF238490	AF238490 Drosophila
26	290.8	5.1	2265	3	AY051880	AY051880 Drosophila
c	280.4	4.9	80146	9	AL583963	AL583963 Human DNA
28	276.4	4.8	2100	3	AF239171	AF239171 Drosophila
29	276.4	4.8	2114	3	AF239171	AF239171 Drosophila
30	262.2	4.6	215989	2	AC102227	AC102227 Mus muscu
31	261.2	4.6	2101	6	AS2135	AS2135 Sequence 1
32	261.2	4.6	2101	6	AR084688	AR084688 Sequence
33	256.6	4.5	3302	10	BC009658	BC009658 Mus muscu
34	254.8	4.5	2160	6	AR170896	AR170896 Sequence
35	249.6	4.4	2305	8	NCCOT1	X97657 N.crassa mR
36	239.2	4.2	245067	2	AC087136	AC087136 Mus muscu
37	232.6	4.1	3018	6	AS2140	AS2140 Sequence 6
38	232.6	4.1	3018	6	AR084691	AR084691 Sequence
39	232.6	4.1	3018	9	HSPROKINX	Z35102 H.sapiens m
40	232.6	4.1	3593	9	BC012085	BC012085 Homo sapi
41	220.6	3.9	2202	8	MCPRTKINA	Z30329 M.crystalli
42	216.6	3.8	1935	6	AR139101	AR139101 Sequence
43	215.8	3.8	1745	3	AF275634	AF275634 Caenorhab
44	215	3.8	4725	9	BC028603	BC028603 Homo sapi
45	215	3.8	4983	6	AX056364	AX056364 Sequence

ALIGNMENTS

RESULT 1
DMU29608
LOCUS
DEFINITION Drosophila melanogaster large tumor suppressor (lats) mRNA, long
5720 bp mRNA linear INV 23-JUL-1995
ACCESSION U29608
VERSION U29608
KEYWORDS transcript, complete cds.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 5720)
AUTHORS Xu,T., Wang,W., Zhang,S., Stewart,R.A. and Yu,W.

Identifying tumor suppressors in genetic mosaics: the *Drosophila* *lats* gene encodes a putative protein kinase
Development 121 (4), 1053-1063 (1995)
95462531
7743921
2. (bases 1 to 5720)
Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.
Direct Submission
Submitted (12-JUN-1995) Tian Xu, Department of Genetics, Yale
University School of Medicine, BCMN Rm. 354, 295 Congress Avenue,
New Haven, CT 06536, USA

FEATURES
SOURCE

hsp70 locus. The *hsp70* gene is located in a complicated genomic region. The *hsp70* gene together with two nearby genes are located in an intron of another gene. The long *hsp70* transcript (5.7kb) is responsible for all *hsp70* activities since the cDNA under the control of the *hsp70* promoter can rescue all *hsp70* mutant phenotypes when it is introduced back into the genome.

gene

exon

even

3

11

11053

exon

exon

exon

COUNT

Db 1021 TGTGCTGCTGTGGAGTGTAGTGTGGAGCATTTCTGTGATGATGATGCTAAATGC 1080
 QY 1081 CACAGGGGCAAGCAGCAGCATCATCCAGCGGCGGAGGAGGAGGCGGTGCGCCCAA 1140
 Db 1081 CACAGGGGCAAGCAGCAGCATCATCCAGCGGCGGAGGAGGAGGCGGTGCGCCCAA 1140
 QY 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGT 1200
 Db 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGT 1200
 QY 1201 ACAAAATACATAGGAATTAATCAGAAATTAACACACCTCTGCGATACACGGGACCAACGG 1260
 Db 1201 ACAAAATACATAGGAATTAATCAGAAATTAACACACCTCTGCGATACACGGGACCAACGG 1260
 QY 1261 AGCAACGATGACATTTACTCTGACTATCACACGCGCAAGCAGCGGATGAGCGCCACC 1320
 Db 1261 AGCAACGATGACATTTACTCTGACTATCACACGCGCAAGCAGCGGATGAGCGCCACC 1320
 QY 1321 CTCGCGCTCTCTGCTCGGAGGTGTATACCGCGCGCGCGCGCGCATTTGTAGTCAAGC 1380
 Db 1321 CTCGCGCTCTCTGCTCGGAGGTGTATACCGCGCGCGCGCGCGCATTTGTAGTCAAGC 1380
 QY 1381 CGAGCGGCTCCATATCCGATTCGCTGCGGTGCGGTGAGTGTGGGTGCGGCAACGG 1440
 Db 1381 CGAGCGGCTCCATATCCGATTCGCTGCGGTGCGGTGAGTGTGGGTGCGGCAACGG 1440
 QY 1441 ACGTGTSCCAAGATGATCAGCGCCCTTAATGCCAAACAACTGATCCGGAAGCCGAGCAT 1500
 Db 1441 ACGTGTSCCAAGATGATCAGCGCCCTTAATGCCAAACAACTGATCCGGAAGCCGAGCAT 1500
 QY 1501 CGAAGCGGACAGCGGAGCAGTCACTACCTGCGCTGCGGTGCGGTGCTGAGTCTCGGAGC 1560
 Db 1501 CGAAGCGGACAGCGGAGCAGTCACTACCTGCGCTGCGGTGCGGTGCTGAGTCTCGGAGC 1560
 QY 1561 CGGTAGTCCGATCGGACAGCCCGCATTCGACACACACCCACCGCGGCTCGAGGAC 1620
 Db 1561 CGGTAGTCCGATCGGACAGCCCGCATTCGACACACACCCACCGCGGCTCGAGGAC 1620
 QY 1621 GGTGGTAAATCAGGTGGAATGTTGATTTCTCGTCCGCAAGCGGTTTCAGTGAAGT 1680
 Db 1621 GGTGGTAAATCAGGTGGAATGTTGATTTCTCGTCCGCAAGCGGTTTCAGTGAAGT 1680
 QY 1681 GGCTCCAGCGCGCGCGCCACGCAATCCACCGCTSCAGCGGCGGCGCACGCCCCACC 1740
 Db 1681 GGCTCCAGCGCGCGCGCCACGCAATCCACCGCTSCAGCGGCGGCGCACGCCCCACC 1740
 QY 1741 GCCAGTGGCGCCAGCAGCGGTAGTGAAGCGGCGATCACCGGCGCTGAACAACGG 1800
 Db 1741 GCCAGTGGCGCCAGCAGCGGTAGTGAAGCGGCGATCACCGGCGCTGAACAACGG 1800
 QY 1801 CCGCGCGGATAGCGCCACCCACTCAGCGMGCAACTCACTGTAAATAACCCARAACGG 1860
 Db 1801 CCGCGCGGATAGCGCCACCCACTCAGCGMGCAACTCACTGTAAATAACCCARAACGG 1860
 QY 1861 GGTGAAGAACCGCAGCAGCTTGACGACAGCTGAAGTCCCTGAACCTATACCCAGG 1920
 Db 1861 GGTGAAGAACCGCAGCAGCTTGACGACAGCTGAAGTCCCTGAACCTATACCCAGG 1920
 QY 1921 CGGAGGAGTGAAGAGTGTGAGGAGCCAGCCGCGGCTTACCTAATTAAGGCGGAGCGG 1980
 Db 1921 CGGAGGAGTGAAGAGTGTGAGGAGCCAGCCGCGGCTTACCTAATTAAGGCGGAGCGG 1980
 QY 1981 AGGAGCAGCAGCGCGCGCCACACCCAGTTACAGGCTTCCATGCTGAGTCCGCGAGTC 2040
 Db 1981 AGGAGCAGCAGCGCGCGCCACACCCAGTTACAGGCTTCCATGCTGAGTCCGCGAGTC 2040
 QY 2041 GCGCACAAATCCCAACATCGGACTACAGGAAATCCCGGAGCAGTGGGATATCTCGG 2100
 Db 2041 GCGCACAAATCCCAACATCGGACTACAGGAAATCCCGGAGCAGTGGGATATCTCGG 2100
 QY 2101 CACTTGGGCGCTCGCGAGCGCCCATACTGTGCTGCTGCGCGCGCGCGCTGGCGAA 2160
 Db 2101 CACTTGGGCGCTCGCGAGCGCCCATACTGTGCTGCTGCGCGCGCGCGCTGGCGAA 2160

Db 2101 CACTTGGGCGCTCGCGAGCGCCCATAACTGTGTGCTGCTGCGCGCGCGCGCTGGCGAA 2160
 QY 2161 GCACAAACACAGTCTTACAGGCCAGGAGTCAAGAGCGGATCATCATCAGAGTGTGAA 2220
 Db 2161 GCACAAACACAGTCTTACAGGCCAGGAGTCAAGAGCGGATCATCATCAGAGTGTGAA 2220
 QY 2221 GAGCAGCAGGTCTCAAAGCCGCTGCTGCAAAACAGCAGTGGGCGGCAATGCGCATCGAG 2280
 Db 2221 GAGCAGCAGGTCTCAAAGCCGCTGCTGCAAAACAGCAGTGGGCGGCAATGCGCATCGAG 2280
 QY 2281 TGCCTGCGGCGAGATTAACAGTTCACAGTTCACCTGTGCGCGCTTCCACCTCTTACCTCAGAA 2340
 Db 2281 TGCCTGCGGCGAGATTAACAGTTCACAGTTCACCTGTGCGCGCTTCCACCTCTTACCTCAGAA 2340
 QY 2341 GTCGCGGCGAGTGTGTCAGCAGCAGCAACAGCAGCAGCGGCGGCGCCACAGCAGCA 2400
 Db 2341 GTCGCGGCGAGTGTGTCAGCAGCAGCAACAGCAGCAGCGGCGGCGCCACAGCAGCA 2400
 QY 2401 TCAGCAGCAGCAATCCAAACCCACCAACGCGCCCTTGGTGGGTCTGAAYAG 2460
 Db 2401 TCAGCAGCAGCAATCCAAACCCACCAACGCGCCCTTGGTGGGTCTGAAYAG 2460
 QY 2461 CAAGCCCAATTCGCTGGAGCCACCGTCTTATGCCAAGAGCATGCGAGCGCGCCAC 2520
 Db 2461 CAAGCCCAATTCGCTGGAGCCACCGTCTTATGCCAAGAGCATGCGAGCGCGCCAC 2520
 QY 2521 GGTGTARAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
 Db 2521 GGTGTARAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
 QY 2581 ACAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2640
 Db 2581 ACAGCAGCAGCAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2640
 QY 2641 TCAGAGGAGCGGATCAACGGGAGCGGAGCGGATGAGCAGAGCTGCGGCAACGGGAA 2700
 Db 2641 TCAGAGGAGCGGATCAACGGGAGCGGAGCGGATGAGCAGAGCTGCGGCAACGGGAA 2700
 QY 2701 TCTCGCGCGAGTGTCTTCCGCGCGCGCTTATCAGAGCAACAAACAGCAGCA 2760
 Db 2701 TCTCGCGCGAGTGTCTTCCGCGCGCGCTTATCAGAGCAACAAACAGCAGCA 2760
 QY 2761 GATCAAAACCGCGAGTGTCAACAAACAAACATACAGATAAAGCAACAGCAACCTGGCGAC 2820
 Db 2761 GATCAAAACCGCGAGTGTCAACAAACAAACATACAGATAAAGCAACAGCAACCTGGCGAC 2820
 QY 2821 GACACACCCATTCGCGCTGCGGCAATACAAATTAACAACTCTCCACAGCGGCGGAATAG 2880
 Db 2821 GACACACCCATTCGCGCTGCGGCAATACAAATTAACAACTCTCCACAGCGGCGGAATAG 2880
 QY 2881 CTCGCGCGCAGCAACGGATCCACCGGACACCGCTCTCTGCGAGCAGTGAAGAA 2940
 Db 2881 CTCGCGCGCAGCAACGGATCCACCGGACACCGCTCTCTGCGAGCAGTGAAGAA 2940
 QY 2941 GATCAAGCAGCCTCGCCATCCCGAGCGGCAAGAGATCTCCAAGGAGGAGGAGGA 3000
 Db 2941 GATCAAGCAGCCTCGCCATCCCGAGCGGCAAGAGATCTCCAAGGAGGAGGAGGA 3000
 QY 3001 GCGCAAGGAGTTCGCGCATCAGCAGTACTCGCGGCAAGCGCTTCAAGTCTTCTCATGGAGCA 3060
 Db 3001 GCGCAAGGAGTTCGCGCATCAGCAGTACTCGCGGCAAGCGCTTCAAGTCTTCTCATGGAGCA 3060
 QY 3061 GCACATAGAGAACGTGATCAAGTCTGCTGCGCGCAGCAGTATCGCAAGAACTCAGTGA 3120
 Db 3061 GCACATAGAGAACGTGATCAAGTCTGCTGCGCGCAGCAGTATCGCAAGAACTCAGTGA 3120
 QY 3121 GAAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCCAAAATCGAGATGAGGAAATGCT 3180
 Db 3121 GAAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCCAAAATCGAGATGAGGAAATGCT 3180
 QY 3181 GAACCAAAAGGAGAGCACTACATTCGATTAAGCGCGGCGCAAGATGGACAGAGCATGTT 3240
 Db 3181 GAACCAAAAGGAGAGCACTACATTCGATTAAGCGCGGCGCAAGATGGACAGAGCATGTT 3240

QY 3241 CGTCAAACTGAAGCCATTGGAGTGGGTGCATTTGGCGAGGTAAACGCTGGTGAACAAAT 3300
 Db 3241 CGTCAAACTGAAGCCATTGGAGTGGGTGCATTTGGCGAGGTAAACGCTGGTGAACAAAT 3300
 QY 3301 CGATACCTCGAACCAATTTGTATGGATGAACAAACCTCGGGAAGGGAGCTTCAAGCG 3360
 Db 3301 CGATACCTCGAACCAATTTGTATGGATGAACAAACCTCGGGAAGGGAGCTTCAAGCG 3360
 QY 3361 GAATCAGGTGCGACAGCTGAAGCGGAGGAGATATCTCGCGAAGCGGACAAATACG 3420
 Db 3361 GAATCAGGTGCGACAGCTGAAGCGGAGGAGATATCTCGCGAAGCGGACAAATACG 3420
 QY 3421 GGTGGTGAAGTTGACTACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGAGGACTA 3480
 Db 3421 GGTGGTGAAGTTGACTACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGAGGACTA 3480
 QY 3481 CATACCAAGTGGTGTATGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGACT 3540
 Db 3481 CATACCAAGTGGTGTATGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGACT 3540
 QY 3541 GGCAGATTCTACATCGCGAGGTCACCTGCGCGTGGACAGCTTCACAAATGGGCTT 3600
 Db 3541 GGCAGATTCTACATCGCGAGGTCACCTGCGCGTGGACAGCTTCACAAATGGGCTT 3600
 QY 3601 CATTACAGAGACATCAAGCTGACACATACTCATCGATAGGACGACACATAAGCT 3660
 Db 3601 CATTACAGAGACATCAAGCTGACACATACTCATCGATAGGACGACACATAAGCT 3660
 QY 3661 CACGACTTTGGCTGTGACGGGATTCGGATGACGACCAACTCGAAGTACTACAGGA 3720
 Db 3661 CACGACTTTGGCTGTGACGGGATTCGGATGACGACCAACTCGAAGTACTACAGGA 3720
 QY 3721 GAACGGAATCACTCGCGCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
 Db 3721 GAACGGAATCACTCGCGCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
 QY 3781 ACCGAAGCCACCTGCTGGAGGCGGACGATCGCGGATCAACCAAGAGTCTCTGGCCCA 3840
 Db 3781 ACCGAAGCCACCTGCTGGAGGCGGACGATCGCGGATCAACCAAGAGTCTCTGGCCCA 3840
 QY 3841 CTCGCTGGTGGCCACCCGACCTACATAGTCCCGAGGCTGGAGAGTGGGTACAC 3900
 Db 3841 CTCGCTGGTGGCCACCCGACCTACATAGTCCCGAGGCTGGAGAGTGGGTACAC 3900
 QY 3901 GCAGCTGTGGCTACTTGGAGCGTGGCGCTCATCTTAYAGATGCTGTGGTCAAGC 3960
 Db 3901 GCAGCTGTGGCTACTTGGAGCGTGGCGCTCATCTTAYAGATGCTGTGGTCAAGC 3960
 QY 3961 GCGCTTTCTGGCCAAAGTCCGCTGGAACGCAACAAAGGTCTCAACTGGGAGAAAC 4020
 Db 3961 GCGCTTTCTGGCCAAAGTCCGCTGGAACGCAACAAAGGTCTCAACTGGGAGAAAC 4020
 QY 4021 SCTGCATATTCGCGCCGACGCGGAGTATCCGCGAGGCTACGGACTGTAGAGGCT 4080
 Db 4021 SCTGCATATTCGCGCCGACGCGGAGTATCCGCGAGGCTACGGACTGTAGAGGCT 4080
 QY 4081 CTGTGCTGCGCTGACAAAGCGCTGGCAAGAGCTGGACAGCTCAAGAGTCAAGAGCAGCTT 4140
 Db 4081 CTGTGCTGCGCTGACAAAGCGCTGGCAAGAGCTGGACAGCTCAAGAGTCAAGAGCAGCTT 4140
 QY 4141 CTTCAAGGGCATCGACTTTTCGGGACATGCGGAAGAGCAAGAGCGCCCTACATACCGGAAT 4200
 Db 4141 CTTCAAGGGCATCGACTTTTCGGGACATGCGGAAGAGCAAGAGCGCCCTACATACCGGAAT 4200
 QY 4201 CAAGCACCACGACACATCTCAACTTTGATCCGCTGGATCCGGAGAGCTGGCTCGAA 4260
 Db 4201 CAAGCACCACGACACATCTCAACTTTGATCCGCTGGATCCGGAGAGCTGGCTCGAA 4260
 QY 4261 TGACTCCACCATGAGCAGCGGCTGATGTCGACCAAGATGACCGCACTTCCACGCTT 4320
 Db 4261 TGACTCCACCATGAGCAGCGGCTGATGTCGACCAAGATGACCGCACTTCCACGCTT 4320

QY 4321 TTTCCAAATTTACCTTCCTCGCTCCTTTTCGACGACAAGCAGCCGCCGATATGACGGACGA 4380
 Db 4321 TTTCCAAATTTACCTTCCTCGCTCCTTTTCGACGACAAGCAGCCGCCGATATGACGGACGA 4380
 QY 4381 TCAGGCGCGGTTTACGCTGCTGAAATGATGCTCTCCATGTGCCCAACACCAACACCCCGC 4440
 Db 4381 TCAGGCGCGGTTTACGCTGCTGAAATGATGCTCTCCATGTGCCCAACACCAACACCCCGC 4440
 QY 4441 CCCCGAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAAACCATTCAGTGGGCTT 4500
 Db 4441 CCCCGAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAAACCATTCAGTGGGCTT 4500
 QY 4501 GCATTGTTAAAGGAAGCGTGTATAGAACTATCTATATACATATATAAATATA 4560
 Db 4501 GCATTGTTAAAGGAAGCGTGTATAGAACTATCTATATACATATATAAATATA 4560
 QY 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATATAATATATATATAT 4620
 Db 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATATAATATATATATATAT 4620
 QY 4621 AT 4680
 Db 4621 AT 4680
 QY 4681 AGATGAAACGAGAGGAGCGAGCTAGGACCTTCGACCTTTAACTGAACATATATATCTT 4740
 Db 4681 AGATGAAACGAGAGGAGCGAGCTAGGACCTTCGACCTTTAACTGAACATATATATCTT 4740
 QY 4741 GTGCACCTACTCTCCACAAATATATATATATATATATATATATATATATATATAT 4800
 Db 4741 GTGCACCTACTCTCCACAAATATATATATATATATATATATATATATATATATAT 4800
 QY 4801 ACTGGAATGAACTTTCTGGTCTCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 Db 4801 ACTGGAATGAACTTTCTGGTCTCAAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 QY 4861 TTAACCTAAAYGAGACGCGAATTTTACCAACCACTTCCTCTCTCTCTCTCTCTCTCT 4920
 Db 4861 TTAACCTAAAYGAGACGCGAATTTTACCAACCACTTCCTCTCTCTCTCTCTCTCTCT 4920
 QY 4921 GATCGGTGGCGGATTCGAACCTAGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4980
 Db 4921 GATCGGTGGCGGATTCGAACCTAGCAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 4980
 QY 4981 CATTCAAAATTTGAGATTCGAGGTGCGATGGAGAACGAGCAAGCAAGCAAGCAAGCAAG 5040
 Db 4981 CATTCAAAATTTGAGATTCGAGGTGCGATGGAGAACGAGCAAGCAAGCAAGCAAGCAAG 5040
 QY 5041 GGCAGCGATATAAGCGGCTCTTATAAGCTTAATCTAAATCTAAATCTAAATCTAAATCT 5100
 Db 5041 GGCAGCGATATAAGCGGCTCTTATAAGCTTAATCTAAATCTAAATCTAAATCTAAATCT 5100
 QY 5101 YATGATGCTGCTATCCAAATTCGCTATCTACTGCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
 Db 5101 YATGATGCTGCTATCCAAATTCGCTATCTACTGCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
 QY 5161 CCCCCCTCCCATCCAAAGCAAACTTAGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
 Db 5161 CCCCCCTCCCATCCAAAGCAAACTTAGACGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 5220
 QY 5221 GTTAGACCAACTTTGTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
 Db 5221 GTTAGACCAACTTTGTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
 QY 5281 GAAATTCACACCTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 Db 5281 GAAATTCACACCTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 QY 5341 AAACAGTACGAGAAATTTGTAATCTTCTTAATGTAATGTAATGTAATGTAATGTAATGTA 5400
 Db 5341 AAACAGTACGAGAAATTTGTAATCTTCTTAATGTAATGTAATGTAATGTAATGTAATGTA 5400
 QY 5401 AATCTATGCTAGAGTTGTGTAGCGCCCTTAAGATGTTTTTTTAGTTTATAGACCGCTAACCG 5460

|||||
Db 5401 AATCTATGCTAGAGTGTGTAGCGCCTAAGATGTTTTTTAGTTTATAGACGCGTAACGC 5460
Qy 5461 TAATCTAGTTTAAATCCCTAACACTAAGCGAGAGTACAGTACATCGTTTTTTTGTGTC 5520
Db 5461 TAATCTAGTTTAAATCCCTAACACTAAGCGAGAGTACAGTACATCGTTTTTTTGTGTC 5520
Qy 5521 GPAGTTCGTTGGAAATGCTTTAACGGGAACGATTGTTTTTCTCTTTAAATAGCTTCA 5580
Db 5521 GPAGTTCGTTGGAAATGCTTTAACGGGAACGATTGTTTTTCTCTTTAAATAGCTTCA 5580
Qy 5581 GTTTGTATGTCGCTGTTTTTATTTATGACTTATATATAGTCCATCTGAATATCGTGA 5640
Db 5581 GTTTGTATGTCGCTGTTTTTATTTATGACTTATATATAGTCCATCTGAATATCGTGA 5640
Qy 5641 TGGAGCCTATTTTAAATGTGAGATCGAGCTAATTTGAAGGAAATACAAACAACTCTGTGT 5700
Db 5641 TGGAGCCTATTTTAAATGTGAGATCGAGCTAATTTGAAGGAAATACAAACAACTCTGTGT 5700
Qy 5701 GCCTTGGCCAAATAGTTTAC 5720
Db 5701 GCCTTGGCCAAATAGTTTAC 5720

RESULT 2
AR201454
LOCUS AR201454 5720 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6359193.
ACCESSION AR201454
VERSION AR201454.1 GI:20252342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5720)
AUTHORS Xu, F., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 1 19-MAR-2002;
FEATURES Location/Qualifiers
source 1. 5720
BASE COUNT 1684 a 1491 c 1457 g 1088 t
ORIGIN

Query Match 99.8%; Score 5708; DB 6; Length 5720;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 5693; Conservative 26; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCTAGCAGCGGCGAGCAACAAACCCAGCAATTAATTTTACTAAATTAAGCCAAACGC 60
Db 1 ATCTAGCAGCGGCGAGCAACAAACCCAGCAATTAATTTTACTAAATTAAGCCAAACGC 60
Qy 61 GCATCGGAAATCCCTGAAAATGCGATTGAATCGACCGGAAAGTGTGGTTCGCAACGC 120
Db 61 GCATCGGAAATCCCTGAAAATGCGATTGAATCGACCGGAAAGTGTGGTTCGCAACGC 120
Qy 121 GAGTGAATCAAGTGAATACGTCGGCAATATACGCAATTTGCTCAAAAGGCAAGGA 180
Db 121 GAGTGAATCAAGTGAATACGTCGGCAATATACGCAATTTGCTCAAAAGGCAAGGA 180
Qy 181 AAACGGAGAAAAGAGGAAAGCAATTAAGTCCGCTGTGGGAAACGCAAAAGGCGA 240
Db 181 AAACGGAGAAAAGAGGAAAGCAATTAAGTCCGCTGTGGGAAACGCAAAAGGCGA 240
Qy 241 GAACAAAGCGGAAAGCGAGGAAATTCGTTGGAACGTTGGAACGCGGAAGGCGA 300
Db 241 GAACAAAGCGGAAAGCGAGGAAATTCGTTGGAACGTTGGAACGCGGAAGGCGA 300
Qy 301 AGCTCCAAAGTTGGCCGCCCATCGATTTCGTGTAGGATCAATTAAGATTCCGAGTGTGCGA 360
Db 301 AGCTCCAAAGTTGGCCGCCCATCGATTTCGTGTAGGATCAATTAAGATTCCGAGTGTGCGA 360
Qy 361 GAATCGGCTCAATCAAAATTAATACTAATATTTTGGTATTCAGATATTCAAATGGA 420

|||||
Db 361 GAATCGGCTCAAAATCAAAATTAATAATATTTTGGTATTCAGATATTCAAATGGA 420
Qy 421 ATTCAATTCATCCCTGCGGACTTTTATTCGAGTCTGCCAACTATTTTGAATTTGAATGT 480
Db 421 ATTCAATTCATCCCTGCGGACTTTTATTCGAGTCTGCCAACTATTTTGAATTTGAATGT 480
Qy 481 GTGCTGCGGCTGCGGCGAGAAATCTCTGATAAAGAGAGAGAAATAAATCGGAAGAACAA 540
Db 481 GTGCTGCGGCTGCGGCGAGAAATCTCTGATAAAGAGAGAGAAATAAATCGGAAGAACAA 540
Qy 541 AATACAAATACAAATGAAATGCGGGAGAGATTTTATGATGCGCAATGCTGATAG 600
Db 541 AATACAAATACAAATGAAATGCGGGAGAGATTTTATGATGCGCAATGCTGATAG 600
Qy 601 GCGAAGGGGGGTTCTCTTATTAATGCAAAATGTAATGTAATGTAATGTAATGTAATGTA 660
Db 601 GCGAAGGGGGGTTCTCTTATTAATGCAAAATGTAATGTAATGTAATGTAATGTAATGTA 660
Qy 661 GTGGAAGAAATCCCGCGCGAGTGATTAATAATCGGACGACAAACAAAGCAGAGCCTAC 720
Db 661 GTGGAAGAAATCCCGCGCGAGTGATTAATAATCGGACGACAAACAAAGCAGAGCCTAC 720
Qy 721 ACCGCGAGAAAGAGAGCGCAACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780
Db 721 ACCGCGAGAAAGAGAGCGCAACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780
Qy 781 GAGATAATAAAGCATCTTAAACCCGCGCTTAGTTCGTTTGTAGTCTGCCACRATATA 840
Db 781 GAGATAATAAAGCATCTTAAACCCGCGCTTAGTTCGTTTGTAGTCTGCCACRATATA 840
Qy 841 GATATTCAAGGCAAAAGGTGTGTGCGCATCGCAGCAACAAAGCAAGTAAAGCATCTATT 900
Db 841 GATATTCAAGGCAAAAGGTGTGTGCGCATCGCAGCAACAAAGCAAGTAAAGCATCTATT 900
Qy 901 TCATACAAACAAACCAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Db 901 TCATACAAACAAACCAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 960
Qy 961 CAAATCAAAATCCCGCGCGCGCATGTGCCCCAGTGTGTGCGTGTGTGTGTGTGTGTGT 1020
Db 961 CAAATCAAAATCCCGCGCGCGCATGTGCCCCAGTGTGTGCGTGTGTGTGTGTGTGTGT 1020
Qy 1021 TGT 1080
Db 1021 TGT 1080
Qy 1081 CACAGGGGAAAGCAGCAGCATCATGATCCAGCGGGGCAAAAGGGGGGGTGTGCCCCAA 1140
Db 1081 CACAGGGGAAAGCAGCAGCATCATGATCCAGCGGGGCAAAAGGGGGGGTGTGCCCCAA 1140
Qy 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTTAACCCGATTTGAAGT 1200
Db 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTTAACCCGATTTGAAGT 1200
Qy 1201 ACAAATACCAATAGGAATTAATCAGAAATACACACTCTGCGATACACGGCGCAACAGG 1260
Db 1201 ACAAATACCAATAGGAATTAATCAGAAATTAACACACTCTGCGATACACGGCGCAACAGG 1260
Qy 1261 AGGCAAGCATGACATCTCTGACTATCACACGCCAAGCAGCGCATGGAGCCGCCACC 1320
Db 1261 AGGCAAGCATGACATCTCTGACTATCACACGCCAAGCAGCGCATGGAGCCGCCACC 1320
Qy 1321 CTCGCCCTCTCTGCTCGGAGTGTGTATACCCGCCGCCGCCCATTTAGGTACGCC 1380
Db 1321 CTCGCCCTCTCTGCTCGGAGTGTGTATACCCGCCGCCGCCCATTTAGGTACGCC 1380
Qy 1381 CGGAGCCGCTCCATATCCGTTTGTGGGCTTGGAGTGTGGTGTGGGCAACGG 1440
Db 1381 CGGAGCCGCTCCATATCCGTTTGTGGGCTTGGAGTGTGGTGTGGGCAACGG 1440
Qy 1441 AGTGTSCCAAGATGATGAGCGCCCTAATGCCAAACAACTGATCCGGAAGCCGAGCAT 1500
Db 1441 AGTGTSCCAAGATGATGAGCGCCCTAATGCCAAACAACTGATCCGGAAGCCGAGCAT 1500

Db	1441	ACGTGTC	CAAGA	TGATG	CAGCG	CCCTA	ATG	CCAA	AACTG	ATCCG	GAAG	CCGAG	CAAT	1500					
Qy	1501	CGA	CGGAC	AGCG	GAG	CAGT	CTAC	TCT	CGCT	GCAGT	CCGGCT	CTTG	AGCT	CCGGAGC	1560				
Db	1501	CGA	CGGAC	ACG	CGC	GAG	CA	TCT	AC	TCT	CGCT	CTG	CA	TCT	CGGAGC	1560			
Qy	1561	CGT	AGCT	CCG	ATG	CG	AGC	CCCC	CA	TTC	GG	ACCA	CA	CCCC	AC	CGAGC	1620		
Db	1561	CGT	AGCT	CCG	ATG	CG	AGC	CCCC	CA	TTC	GG	ACCA	CA	CCCC	AC	CGAGC	1620		
Qy	1621	GGT	GGT	TAAT	CC	AGT	TG	GAAT	TTC	T	CGT	CG	CA	AGCG	TTT	CAGT	GA	1680	
Db	1621	GGT	GGT	TAAT	CC	AGT	TG	GAAT	TTC	T	CGT	CG	CA	AGCG	TTT	CAGT	GA	1680	
Qy	1681	GGT	CC	ACCG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1740
Db	1681	GGT	CC	ACCG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1740
Qy	1741	GCC	AGT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1800
Db	1741	GCC	AGT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1800
Qy	1801	CCG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1860
Db	1801	CCG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1860
Qy	1861	GCT	GA	AA	AC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1920
Db	1861	GCT	GA	AA	AC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1920
Qy	1921	CGG	AG	CG	AGT	GG	AG	CA	GT	GG	AG	CA	GT	GG	AG	CA	GT	GG	1980
Db	1921	CGG	AG	CG	AGT	GG	AG	CA	GT	GG	AG	CA	GT	GG	AG	CA	GT	GG	1980
Qy	1981	AGG	AG	CA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2040
Db	1981	AGG	AG	CA	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2040
Qy	2041	GCC	CA	CA	CA	AT	CC	CA	CA	AT	CC	CA	CA	AT	CC	CA	CA	AT	2100
Db	2041	GCC	CA	CA	CA	AT	CC	CA	CA	AT	CC	CA	CA	AT	CC	CA	CA	AT	2100
Qy	2101	CAC	CT	GG	GG	GG	CT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2160
Db	2101	CAC	CT	GG	GG	GG	CT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2160
Qy	2161	GCC	CA	AC	CA	CG	AGT	CT	AC	CA	CG	AG	AGT	CT	AC	CA	CG	AG	2220
Db	2161	GCC	CA	AC	CA	CG	AGT	CT	AC	CA	CG	AG	AGT	CT	AC	CA	CG	AG	2220
Qy	2221	GAG	CA	CG	AGT	CC	AA	AG	CC	GT	GT	CG	AA	AG	CC	GT	GT	CG	2280
Db	2221	GAG	CA	CG	AGT	CC	AA	AG	CC	GT	GT	CG	AA	AG	CC	GT	GT	CG	2280
Qy	2281	TGC	CT	CG	CG	CG	CA	GT	CT	AC	CG	CG	CG	CG	CG	CG	CG	CG	2340
Db	2281	TGC	CT	CG	CG	CG	CA	GT	CT	AC	CG	CG	CG	CG	CG	CG	CG	CG	2340
Qy	2341	GT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2400
Db	2341	GT	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	2400
Qy	2401	TC	AG	CA	CG	CA	GT	TC											

QY	2581	ACAGCAGCAGCAGCAGCAGCAACAGCAACTGACGCCCTTGAGGGTGTCTCAGGCACAGGC	2640
DB	2581		
QY	2581	ACAGCAGCAGCAACAGCAGCAACAGCAACTGACGCCCTTGAGGGTGTCTCAGGCACAGGC	2640
DB	2581		
QY	2641	TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCAAACGGAAA	2700
DB	2641		
QY	2701	TCCTGGCGCGCAGATGCTCCGCGCGGCCCTATCAGAGCAACAACAACAACAGCGGA	2760
DB	2701		
QY	2761	GATCAAAACGCGCAGCTGCAACACAAACAACATACAGATTAAGCAACAGCAACTGGCCAC	2820
DB	2761		
QY	2821	GACACCAACCTTCGCGCTGCCAAATAACAACCTCTCAACACAGCGCGCGCAATAG	2880
DB	2821		
QY	2881	CTCGGGCGCAGCAACCGGATCCACGGCACACCGCTCTCTGCACACAGCTGCAAGAA	2940
DB	2881		
QY	2941	GATCAAGCAGCGCTCGCCATCCGCGGCGCAAGAAGATCTCAAGGAGAAAGGAGGAGGA	3000
DB	2941		
QY	3001	GGCAGAGGAGTTCCGCATCAGCAGCAGTACTGCCGCCAAGCCCTCAAGTCTTCATGAGCA	3060
DB	3001		
QY	3061	GCACATAGAGACGTGATCAAGTCTGATCGCCGCGCAGCTATCGCAAGAATCAGCTGGA	3120
DB	3061		
QY	3121	GAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAATGCT	3180
DB	3121		
QY	3181	GAACCAAAAGGAGAGCAACTACATCTGATTGAAGCGCGCAAGATGGACAAGACATGTT	3240
DB	3181		
QY	3241	CGTCAAACTGAAGCCCATTTGAGTGGGTGCATTTGGCAGGTAAAGCTGGTGCACAAAT	3300
DB	3241		
QY	3301	CGATACCTTCGAACCATTTGATCCGATGAAACCCCTGCGGAAGCGGACGTTCTCAAGCG	3360
DB	3301		
QY	3361	GAATCAGTGGCACACGTGAAGCGCAGAGGGATATCTTCGCGGAAGCGGACATAACTG	3420
DB	3361		
QY	3421	GGTGTGAAGTTGATACAGCTTCAGGACAGGATATCTGTACTTTGTGATGGACTA	3480
DB	3421		
QY	3481	CATACCAGTGGTGTATGATGTCGCTCATCAAACTGGGCAATTTTCGAGGAGAACT	3540
DB	3481		
QY	3541	GGCCAGATTCTACATCGCGAGGTCACTGCGCGGTGGACAGCTTCACAAAATGGCTT	3600
DB	3541		
QY	3601	CATTACAGACATCAAGCTTGACAACATACTCATGATAGGACGGACACATAAAGCT	3660
DB	3601		

QY 3661 CACCGACTTTGGCTGTGACGGGATTCGATGGAGCCCAAACTCGAAGTACTACACGGA 3720
 Db 3661 CACCGACTTTGGCTGTGACGGGATTCGATGGAGCCCAAACTCGAAGTACTACACGGA 3720
 QY 3721 GAACGGCAATCACTCCGCGCAGGACTCGATGGAGCCCTGGGAGGAAATACCTCCGAGAACGG 3780
 Db 3721 GAACGGCAATCACTCCGCGCAGGACTCGATGGAGCCCTGGGAGGAAATACCTCCGAGAACGG 3780
 QY 3781 ACCGAAGCCACCGTCTGCGAGAGGCGGCGATGCGCGATCAACAAAGAGTCTCTGGCCCA 3840
 Db 3781 ACCGAAGCCACCGTCTGCGAGAGGCGGCGATGCGCGATCAACAAAGAGTCTCTGGCCCA 3840
 QY 3841 CTCGCTGGTGGGACCCGAACTACATAGCTCCCGAGGCTGCTGGAGAGTGGGTACAC 3900
 Db 3841 CTCGCTGGTGGGACCCGAACTACATAGCTCCCGAGGCTGCTGGAGAGTGGGTACAC 3900
 QY 3901 GCAGCTGTCGGACTACTGAGCGGTGGCGTCACTCTTAYAGATGCTGGTGGGTACGC 3960
 Db 3901 GCAGCTGTCGGACTACTGAGCGGTGGCGTCACTCTTAYAGATGCTGGTGGGTACGC 3960
 QY 3961 GCCCTTTCTGGCCACAGTCCGCTGGAAACGCAACAAAAGGTCACTCAACTGGGAGAAAC 4020
 Db 3961 GCCCTTTCTGGCCACAGTCCGCTGGAAACGCAACAAAAGGTCACTCAACTGGGAGAAAC 4020
 QY 4021 SCTGCATATTCGCGCGCAGCGCGAGTTATCCCGGAGGCTACGACTTGTGATAAGGAGCT 4080
 Db 4021 SCTGCATATTCGCGCGCAGCGCGAGTTATCCCGGAGGCTACGACTTGTGATAAGGAGCT 4080
 QY 4081 CTGTGCGTGGTGCACAGCGCTGGGCAAGAGCGTGGAGCGAGGTCAAGAGCCACGACTT 4140
 Db 4081 CTGTGCGTGGTGCACAGCGCTGGGCAAGAGCGTGGAGCGAGGTCAAGAGCCACGACTT 4140
 QY 4141 CTTCAAGGCGATCGACTTTTGGGACATCGGAGCAGAAAGCGCCCTACATACCGGAAT 4200
 Db 4141 CTTCAAGGCGATCGACTTTTGGGACATCGGAGCAGAAAGCGCCCTACATACCGGAAT 4200
 QY 4201 CAAGCACCCACGAGCAGTTCGACGACATCCAACTTTGATCCCGTGGATCCGAGAACTCG 4260
 Db 4201 CAAGCACCCACGAGCAGTTCGACGACATCCAACTTTGATCCCGTGGATCCGAGAACTCG 4260
 QY 4261 TGACTCCACCATGAGCAGCGGCGATGATGTCGACGAGATGACCGGCTTTCCACGGCTT 4320
 Db 4261 TGACTCCACCATGAGCAGCGGCGATGATGTCGACGAGATGACCGGCTTTCCACGGCTT 4320
 QY 4321 TTTTCAATTTACTCTTCCGCTGCTTCTGACGACAGCAGCGCGCGATGATGACGAGCA 4380
 Db 4321 TTTTCAATTTACTCTTCCGCTGCTTCTGACGACAGCAGCGCGCGATGATGACGAGCA 4380
 QY 4381 TCAGGCGCGGTTTACGCTGAAATGGATGCTCTCCATGTCGCGCAACCAACACCCCGC 4440
 Db 4381 TCAGGCGCGGTTTACGCTGAAATGGATGCTCTCCATGTCGCGCAACCAACACCCCGC 4440
 QY 4441 CCCCGAATCATTTAGTCAATAGTCAAAAAGGGATAGAACCATTTAGTGGGCTT 4500
 Db 4441 CCCCGAATCATTTAGTCAATAGTCAAAAAGGGATAGAACCATTTAGTGGGCTT 4500
 QY 4501 GCATTTGAAGGAAGCGTGGCTATAGAATGAACATCTATATACATTTATAAATTTATA 4560
 Db 4501 GCATTTGAAGGAAGCGTGGCTATAGAATGAACATCTATATACATTTATAAATTTATA 4560
 QY 4561 GGAGCAGTGAAGGGGAGCTACGCTATACATACAAATATATACATATATTTGATAT 4620
 Db 4561 GGAGCAGTGAAGGGGAGCTACGCTATACATACAAATATATACATATATTTGATAT 4620
 QY 4621 AT 4680
 Db 4621 AT 4680
 QY 4681 AGATGAACGAGAGGAGCGAGTCCGACCTTCGACCTTTTAACTGACATAGTATATCTT 4740
 Db 4681 AGATGAACGAGAGGAGCGAGTCCGACCTTCGACCTTTTAACTGACATAGTATATCTT 4740
 QY 4741 GTGCACACTACTCCACAAATATATATATATATATATATATATATATATATATATAT 4800

Db 4741 GTGCACACTACTCCACAAATATATATATATATATATATATATATATATATATATAT 4800
 QY 4801 ACTGGAAATCGAACCTTTCTGTGTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 Db 4801 ACTGGAAATCGAACCTTTCTGTGTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 QY 4861 TTAACACTAAAGAGAGCGGAAATTTACCAACACACTTCACTCTCTCTCTCTCTCTCTCT 4920
 Db 4861 TTAACACTAAAGAGAGCGGAAATTTACCAACACACTTCACTCTCTCTCTCTCTCTCTCT 4920
 QY 4921 GATCGGTGCGCGGATTCGAACCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4980
 Db 4921 GATCGGTGCGCGGATTCGAACCTTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4980
 QY 4981 CATTCAGAAATTCAGATTCGAGGTGCGATGGAGCAAGCAAGCAAGCAAGCAAGCAAG 5040
 Db 4981 CATTCAGAAATTCAGATTCGAGGTGCGATGGAGCAAGCAAGCAAGCAAGCAAGCAAG 5040
 QY 5041 GGCAGCGATATAAGCGGGTCTTATAGCCCTAATCTAAATCTAACTGGGAGCAAGCAAG 5100
 Db 5041 GGCAGCGATATAAGCGGGTCTTATAGCCCTAATCTAAATCTAACTGGGAGCAAGCAAG 5100
 QY 5101 YATGATGCTCTGCTATCCAAATTCGCTATCACTGCTCTTCACTGCTGCTGCTGCTGCT 5160
 Db 5101 YATGATGCTCTGCTATCCAAATTCGCTATCACTGCTCTTCACTGCTGCTGCTGCTGCT 5160
 QY 5161 CCCCCCTTCCCTCCCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5220
 Db 5161 CCCCCCTTCCCTCCCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5220
 QY 5221 GTTAGACCAACTTGTGAATGCAATGAAATGTTTACGCCACGAGGAAACGCGGGG 5280
 Db 5221 GTTAGACCAACTTGTGAATGCAATGAAATGTTTACGCCACGAGGAAACGCGGGG 5280
 QY 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGCAAGCAAGCAAGCAAGCAAGCAAG 5340
 Db 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGCAAGCAAGCAAGCAAGCAAGCAAG 5340
 QY 5341 AAACAGTACGAGAAATTCGCTCTTAAATGTAATATGTAAGCAACACGTTAATGCT 5400
 Db 5341 AAACAGTACGAGAAATTCGCTCTTAAATGTAATATGTAAGCAACACGTTAATGCT 5400
 QY 5401 AATCTATGCTAGAGTGTGTAGCGCCTAAGATGTTTTTATGTTTATAGACGCTAACCG 5460
 Db 5401 AATCTATGCTAGAGTGTGTAGCGCCTAAGATGTTTTTATGTTTATAGACGCTAACCG 5460
 QY 5461 TAATCTAGTTTAAATTCCTAACACTAAGCGAGAGTACAGTACATGTTTTTTTGTGTC 5520
 Db 5461 TAATCTAGTTTAAATTCCTAACACTAAGCGAGAGTACAGTACATGTTTTTTTGTGTC 5520
 QY 5521 GTAGGTTCTTGGAAATGCTTAAACGGAAGCAAGTTGTTTTCTCTTAAATAGCTTCA 5580
 Db 5521 GTAGGTTCTTGGAAATGCTTAAACGGAAGCAAGTTGTTTTCTCTTAAATAGCTTCA 5580
 QY 5581 GTTGTATGCTGCTGTTTTTATATGACTTATATATATATATATATATATATATATATAT 5640
 Db 5581 GTTGTATGCTGCTGTTTTTATATGACTTATATATATATATATATATATATATATATAT 5640
 QY 5641 TGGAGCCTATTTAAATGTGAGATCGAGTAAATGAAGGAAATACAAACAACTCTGTGT 5700
 Db 5641 TGGAGCCTATTTAAATGTGAGATCGAGTAAATGAAGGAAATACAAACAACTCTGTGT 5700
 QY 5701 GCCTTGGCCCAATAGTTTAC 5720
 Db 5701 GCCTTGGCCCAATAGTTTAC 5720

RESULT 3

DROWARTS

LOCUS

DEFINITION

Drosophila melanogaster tumor suppressor (warts) mRNA linear INV 03-AUG-1995 complete cds. 5360 bp

us-09-763-334-7.rge

[illegible]

QY	3119	GAGAAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAATG	3178
Db	3153	GAGAAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAATG	3212
QY	3179	CTGAACCAAAAGGAGACAACACTACATTTCGATTGAAGCGGCCCAAGATGGACAAGAGCATG	3238
Db	3213	CTGAACCAAAAGGAGAGCAACTACATTTCGATTGAAGCGGCCCAAGATGGACAAGAGCATG	3272
QY	3239	TTTCGTCAAACCTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAAACGCTGGTGAGCAAA	3298
Db	3273	TTTCGTCAAACCTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAAACGCTGGTGAGCAAA	3332
QY	3299	ATCGATACCTCGAACCACTTTGTATGCGATGAANAACCTCGGNAAGCGGAGTCTCTCAAG	3358
Db	3333	ATCGATACCTCGAACCACTTTGTATGCGATGAANAACCTCGGNAAGCGGAGTCTCTCAAG	3392
QY	3359	CGGAATCAGGTTGGCACACGTTGAAGCGCCGAGAGGGATATCCTCGCGAAGCCGACAAATAAC	3418
Db	3393	CGGAATCAGGTTGGCACACGTTGAAGCGCCGAGAGGGATATCCTCGCGAAGCCGACAAATAAC	3452
QY	3419	TGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAGATAAATCTGTACTTTTGTGATGAC	3478
Db	3453	TGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAGATAAATCTGTACTTTTGTGATGAC	3512
QY	3479	TACATACCAAGTGGTGTCTCATCTGCTGCTCATCAACTGGGCATTTTTCGAGGAGGAA	3538
Db	3513	TACATACCAAGTGGTGTCTCATCTGCTGCTCATCAACTGGGCATTTTTCGAGGAGGAA	3572
QY	3539	CTGGCCAGATTCTACATCGCCGAGGTACCTGGCCGCTGGACAGCGTTACAAATATGGCC	3598
Db	3573	CTGGCCAGATTCTACATCGCCGAGGTACCTGGCCGCTGGACAGCGTTACAAATATGGCC	3632
QY	3599	TTCAATTCACAGAGACATCAAGCTTGACAAACATCTCATAGGACGACGACACATAAAG	3658
Db	3633	TTCAATTCACAGAGACATCAAGCTTGACAAACATCTCATAGGACGACGACACATAAAG	3692
QY	3659	CTCACCAGCTTTGGCCTGTGCACGGGATTCGGATGGACGACCAACTCGAAGTACTACGAG	3718
Db	3693	CTCACCAGCTTTGGCCTGTGCACGGGATTCGGATGGACGACCAACTCGAAGTACTACGAG	3752
QY	3719	GAGAACGGCAATCACTCGGCCGAGGACTCGATGAGCGCCCTGGAGGAATACTCCGAGAAC	3778
Db	3753	GAGAACGGCAATCACTCGGCCGAGGACTCGATGAGCGCCCTGGAGGAATACTCCGAGAAC	3812
QY	3779	GGACCGAAGCCACCGCTGTGTGAGAGGCGACGGATGCGCGATCACCAAGAGTCTTGGCC	3838
Db	3813	GGACCGAAGCCACCGCTGTGTGAGAGGCGACGGATGCGCGATCACCAAGAGTCTTGGCC	3872
QY	3839	CATCTCGTGTGGGCAACCCGAACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTAC	3898
Db	3873	CATCTCGTGTGGGCAACCCGAACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTAC	3932
QY	3899	ACGACGCTGTGCGACTACTGGAGCTGGCGCTCATCTCTTAYGAGATGCTTGGTGGGTCAG	3958
Db	3933	ACGACGCTGTGCGACTACTGGAGCTGGCGCTCATCTCTTAYGAGATGCTTGGTGGGTCAG	3992
QY	3959	CCGCGCTTTCTGCGCAACAGTCCGCTGGAAACGCAACAAAGGTCTCACTCGGAGAAA	4018
Db	3993	CCGCGCTTTCTGCGCAACAGTCCGCTGGAAACGCAACAAAGGTCTCACTCGGAGAAA	4052
QY	4019	ACSTGCAATATCCGCGCAGCCGAGTTATCCGCGAGGCTACGGACTTGTATAAGGAGG	4078
Db	4053	ACSTGCAATATCCGCGCAGCCGAGTTATCCGCGAGGCTACGGACTTGTATAAGGAGG	4112
QY	4079	CTCTGTGCGTGGCTGACAAGCGGCTGGGCAAGAGCGTTGGACGAGTCAAGGCCACGAC	4138
Db	4113	CTCTGTGCGTGGCTGACAAGCGGCTGGGCAAGAGCGTTGGACGAGTCAAGGCCACGAC	4172
QY	4139	TTCTTCAAGGCGATCGACTTTTGGGACATCGCGGAAGCAGAAGCGCCCTACATACCGGAA	4198
Db	4173	TTCTTCAAGGCGATCGACTTTTGGGACATCGCGGAAGCAGAAGCGCCCTACATACCGGAA	4232

Qy	4199	ATCAAGCACCCACACGACACATCCAACTTTGATCCCGTGGATCCGGAGAGCTTCGCGTCG	4258
Db	4233	ATCAAGCACCCACACGACACATCCAACTTTGATCCCGTGGATCCGGAGAGCTTCGCGTCG	4292
Qy	4259	AATGACTCCACCATGAGCAGCGGGGATGATCTGCACACAGAAATGACCGCACATTCCTCCACGGC	4318
Db	4293	AATGACTCCACCATGAGCAGCGGGGATGATCTGCACACAGAAATGACCGCACATTCCTCCACGGC	4352
Qy	4319	TTTTTTCGAATTTACCTTCCGTCGCTTCTTCGACACAAAGCAGCGCCGGATATGACGGAC	4378
Db	4353	TTTTTTCGAATTTACCTTCCGTCGCTTCTTCGACACAAAGCAGCGCCGGATATGACGGAC	4412
Qy	4379	GATCAGCGCGCGTTTACGTCGTGAATGGATGCTCTCCATGTCGCCAACACCAACACCCC	4438
Db	4413	GATCAGCGCGCGTTTACGTCGTGAATGGATGCTCTCCATGTCGCCAACACCAACACCCC	4472
Qy	4439	GCCCCGGAATCATTTGTTAGTCAAAATAGTCACAAAAGGGGATAGAAAACCATTTGAGTGGGC	4498
Db	4473	GCCCCGGAATCATTTGTTAGTCAAAATAGTCACAAAAGGGGATAGAAAACCATTTGAGTGGGC	4532
Qy	4499	TTGCATTTGAAGGAAGCGTGGCTATAGAATGAACATATCTATATACATTTATAAATTA	4558
Db	4533	TTGCATTTGAAGGAAGCGTGGCTATAGAATGAACATATCTATATACATTTATAAATTA	4592
Qy	4559	TAGGAGACAGTAGAGCGGGGAGCTACGTATATACATACAAAATATATACATATATTTGAT	4618
Db	4593	TAGGAGACAGTAGAGCGGGGAGCTACGTATATACATACAAAATATATACATATATTTGAT	4652
Qy	4619	ATPATATATATATATATGTCGCGTAGGGCATGAAGTGAATTAATATAAACGGAGCCGAGT	4678
Db	4653	ATPATATATATATATATGTCGCGTAGGGCATGAAGTGAATTAATATAAACGGAGCCGAGT	4712
Qy	4679	AGAGATGAACGAGAGGAGCGAGTCAGGACCTTCGACCTTTAACTGAACTAGATATATCC	4738
Db	4713	AGAGATGAACGAGAGGAGCGAGTCAGGACCTTCGACCTTTAACTGAACTAGATATATCC	4772
Qy	4739	TTGTGCACCTACTACTCCACACAAATATATATTTTAAATTTGTTAGAATTCAAAGGRC	4798
Db	4773	TTGTGCACCTACTACTCCACACAAATATATATTTTAAATTTGTTAGAATTCAAAGGRC	4832
Qy	4799	CAACTGGAATCGAACCTTCTCGTGCCTCAAAGCAAGCAAAAGCAAAACAAAACG	4858
Db	4833	CAACTGGAATCGAACCTTCTCGTGCCTCAAAGCAAGCAAAAGCAAAACAAAACG	4892
Qy	4859	CCCTTAACTTAAATGAGACGCGAATTAACCAACCACTTCACCTCTCTCTTCTCCACT	4918
Db	4893	CCCTTAACTTAAATGAGACGCGAATTAACCAACCACTTCACCTCTCTCTTCTCCACT	4952
Qy	4919	CCGATCGGTGGCGGATTCGAACTCAGCAGCGCTTGCATCCGCGCATCCCATTKACTT	4978
Db	4953	CCGATCGGTGGCGGATTCGAACTCAGCAGCGCTTGCATCCGCGCATCCCATTKACTT	5012
Qy	4979	CCCATTCAGAAATGAGATTCCGAGGTGTCGATGAGAAACGAGACCAAAAAGTCGC	5038
Db	5013	CCCATTCAGAAATGAGATTGCGAGGTGTCGATGAGAAACGAGACCAAAAAGTCGC	5072
Qy	5039	ACGGACGATATAACGGGTCTTATAAGCCCTAATCTAACTCTAACTGGGAGACAGGA	5098
Db	5073	ACGGACGATATAACGGGTCTTATAAGCCCTAATCTAACTCTAACTGGGAGACAGGA	5132
Qy	5099	CCVATGTA-----TGTCCTGCTATCCAAATTCGCTATCACTGCTCTTCAMCTGW	5147
Db	5133	CCVATGTAATGTTATTAGTGCTCTGCTATCCAAATTCGCTATCACTGCTCTTCATCTGT	5192
Qy	5148	GTACGACC--CCACCCCGCCCTCCCTCCCAATCCAAAGACAAACTTAGACGTAGCCCTATG	5205
Db	5193	GTACCCCGCTCCCGCCACACCCCTCCCAATCCCAAGAAACAACTTAGACGTAGCCCTATG	5252
Qy	5206	TGAAAAGCTAGCAATGTTAGACCAACTTGTGTAATGCCAAATGAAATGTTTAGCCCCAC	5265
Db	5253	TGAAAAGCTAGCAATGTTAGACCAACTTGTGTAATGCCAAATGAAATGTTTAGCCCCCG	5312
Qy	5266	GAGGAAACCGCGGGGAAATTCACACTATTTCTCTGTATAGCAACGG	5313

Db 106178 GTGCTGCAACACAGCAGTGGGCGCCCAATCGCCATCGAGTGGCTCGGCCAGCAATTCACCA 106119
 Qy 2303 GTCCACGCTGTGGCGCTTCCACCTCTTACCTCAGAAAGTCCGCGGAGTGGTGCAGCAG 2362
 Db 106118 GTCCACGCTGTGGCGCTTCCACCTCTTACCTCAGAAAGTCCGCGGAGTGGTGCAGCAG 106059
 Qy 2363 CAGCAACAGGAGCAGCGGCGGCGCCACAGCAGCAGATCAGCAACAGCAATCCAAACCA 2422
 Db 106058 CAGCAACAGGAGCAGCGGCGGCGCCACAGCAGCAGATCAGCAACAGCAATCCAAACCA 105999
 Qy 2423 CCAAGCGCMACCAACACCGCCCTTGGTGGTCTGAAYAGCAAGCCCAATTCGCTGGAGCCA 2482
 Db 105998 CCAAGCGCCCAACACCGCCCTTGGTGGTCTGAATAGCAAGCCCAATTCGCTGGAGCCA 105939
 Qy 2483 CCGTCTCTATGCCAAGAGCATCAGCGCCAAAGCGGCCACGGTGGTARAGCAGCAGCAA 2539
 Db 105938 CCGTCTCTATGCCAAGAGCATCAGCGCCAAAGCGGCCACGGTGGTARAGCAGCAGCAA 105879
 Qy 2540 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2599
 Db 105878 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 105819
 Qy 2600 CAACAGCAACTGCAGGCTTTCAGGCTGCTCCAGCAGCAGGCTCAGA 2645
 Db 105818 CAACAGCAACTGCAGGCTTTCAGGCTGCTCCAGCAGCAGGCTCAGAGGAGCGGGATCAA 105759
 Qy 2646 -GGGAGCGGGATCAACGGGAGCGGAAACGGGATCAGCAGAACTGGCCAAACGGGAAATCCT 2704
 Db 105758 CGGAGCGGGATCAACGGGAGCGGAAACGGGATCAGCAGAACTGGCCAAACGGGAAATCCT 105699
 Qy 2705 GCGCGGAGATGCTTCGCGCGCGCCCTATCAGAGCAACAAACAAACAGCGAGATC 2764
 Db 105698 GCGCGGAGATGCTTCGCGCGCGCCCTATCAGAGCAACAAACAAACAGCGAGATC 105639
 Qy 2765 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCACAGCAACCTGGCGACGACA 2824
 Db 105638 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCACAGCAACCTGGCGACGACA 105579
 Qy 2825 CCACCCATTCGCGCTGCAACAAACAAACATACAGATAGCACAGCAACCTGGCGACGACA 2884
 Db 105578 CCACCCATTCGCGCTGCAACAAACAAACATACAGATAGCACAGCAACCTGGCGACGACA 105519
 Qy 2885 GCGCGGAGCAGATCCACCGGACACACCGCTCTCTGTCGACAGTGTGCAAGAAATC 2944
 Db 105518 GCGCGGAGCAGATCCACCGGACACACCGCTCTCTGTCGACAGTGTGCAAGAAATC 105459
 Qy 2945 AAGCACGCTCGCCATCCCGGAGCGGAAGATCTCCAGGAGAGGAGGAGGAGCGC 3004
 Db 105458 AAGCACGCTCGCCATCCCGGAGCGGAAGATCTCCAGGAGAGGAGGAGGAGCGC 105399
 Qy 3005 AAGGATTCGCGCATCAGCAGTACTCGCGCAAGCCTTCAAGTTCTTCATGGAGCAGCAC 3064
 Db 105398 AAGGATTCGCGCATCAGCAGTACTCGCGCAAGCCTTCAAGTTCTTCATGGAGCAGCAC 105339
 Qy 3065 ATAGAGACGTGATCAGTGTATCGCGAGCGCAGTATCGCAAGATCAGCTGGAGAG 3124
 Db 105338 ATAGAGACGTGATCAGTGTATCGCGAGCGCAGTATCGCAAGATCAGCTGGAGAG 105279
 Qy 3125 GAGATGCACAAGTGGGACTCCCGGATCAGACCCAAATCGAGATGAGGAAATCTGAAC 3184
 Db 105278 GAGATGCACAAGTGGGACTCCCGGATCAGACCCAAATCGAGATGAGGAAATCTGAAC 105219
 Qy 3185 CAAAGGAGAGCAACTACATTCGATTAAGCGCGCCCAAGATGGAACAGAGCATGTCGTC 3244
 Db 105218 CAAAGGAGAGCAACTACATTCGATTAAGCGCGCCCAAGATGGAACAGAGCATGTCGTC 105159
 Qy 3245 AAACCTGAAGCCATTGAGTGGGTGCAATTTGGCGAGGTAAAGCTGGTGAAGAAATCGAT 3304
 Db 105158 AAACCTGAAGCCATTGAGTGGGTGCAATTTGGCGAGGTAAAGCTGGTGAAGAAATCGAT 105099
 Qy 3305 ACCTTCGAACCAATTTGATGCGATCAAAACCTTCGGAAGAGCGGAGCTTCTCAAGCGGAAT 3364
 Db 105098 ACCTTCGAACCAATTTGATGCGATCAAAACCTTCGGAAGAGCGGAGCTTCTCAAGCGGAAT 105039

Qy 3365 CAGGTGCGACACGTGAAGCGCGAGAGGATATCTCCGGAAGCCGACAAATAACTGGGTG 3424
 Db 105038 CAGGTGCGACACGTGAAGCGCGAGAGGATATCTCCGGAAGCCGACAAATAACTGGGTG 104979
 Qy 3425 GTGAAGTTGTACTACAGTTCACAGGACAAAGATATCTGTACTTTGTGATGGACTACATA 3484
 Db 104978 GTGAAGTTGTACTACAGTTCACAGGACAAAGATATCTGTACTTTGTGATGGACTACATA 104919
 Qy 3485 C----- 3485
 Db 104918 CCAGGTGAGTGGAAACTTAAGCACTCCAAAGATGTGTATCGGAGTGCAAATGTCTATGTG 104859
 Qy 3486 -----CAGTGGTGTATGTATGTCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 3536
 Db 104858 GATTGTCGACAGGTGGTGTATGTATGTCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 104799
 Qy 3537 AACTGGCCAGATTCATACATCCCGAGGTCACTCCGCGGTGGAGAGGCTTCACAAATGG 3596
 Db 104798 AACTGGCCAGATTCATACATCCCGAGGTCACTCCGCGGTGGAGAGGCTTCACAAATGG 104739
 Qy 3597 GCTTCATT----- 3604
 Db 104738 GCTTCATTACAGGTAATTAGCAGTTGCTTCAGTTATACAGTTCACATTTACATATTGTT 104679
 Qy 3605 -----CACAGACATCAAGCTTGACAACTACTCATCATGATAGGACCGGACACA 3653
 Db 104678 ATGGATTATCCACAGAGACATCAAGCTGACAACTACTCATCATGATAGGACCGGACACA 104619
 Qy 3654 TAAAGCTCAGGACTTTGGCTGTGCAGGGATTCGATGGACGACCAACTCGAGTACT 3713
 Db 104618 TAAAGCTCAGGACTTTGGCTGTGCAGGGATTCGATGGACGACCAACTCGAGTACT 104559
 Qy 3714 ACCAGGAGAAC----- 3724
 Db 104558 ACCAGGAGAACGTTGAGATGGAACCCAGCTAACGTAAGTGTTCACAGCTTTAATAATAT 104499
 Qy 3725 -----GGCAATCACTCGCGCAGGACTTCGATGGAGCCCTGGGAGGAAATA 3768
 Db 104498 CATTCCTACTCTGATAGCAATCACTCGCGCAGGACTTCGATGGAGCCCTGGGAGGAAATA 104439
 Qy 3769 CTCCGAGAACGGACCGGAGCCACCGCTGCTGGAG----- 3802
 Db 104438 CTCCGAGAACGGACCGGAGCCACCGCTGCTGGAGAGTGTGAGTTTCGTACACTTAATTAAT 104379
 Qy 3803 -----AGCGCAGGATCGCGGATCACCACAA 3827
 Db 104378 AGGATCGGCCCTTAATAAAGTATGCCAACTCCGAGCGGACGATGCGGATCACCACAA 104319
 Qy 3828 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAATCATATAGTCCCGAGGAGTGTGAGAG 3887
 Db 104318 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAATCATATAGTCCCGAGGAGTGTGAGAG 104259
 Qy 3888 GAGTGGGTACACGACGCTGCGGACTTCTGGCCAACTGCGCTGGAAACGACAAACAA 3947
 Db 104258 GAGTGGGTACACGACGCTGCGGACTTCTGGCCAACTGCGCTGGAAACGACAAACAA 104199
 Qy 3948 TGGTGGGTGAGCGCCCTTCTGGCCAACTGCGCTGGAAACGACAAACAA 3998
 Db 104198 TGGTGGGTGAGCGCCCTTCTGGCCAACTGCGCTGGAAACGACAAACAA 104139
 Qy 3999 -----AGGTCA 4004
 Db 104138 CTTTGGATTCTTATAAGCCAACTGATTAACAACATGACTTCTTCATTACAGGTCA 104079
 Qy 4005 TCAACTGGGAAACACTGCATATTCGCGCAGGCGGAGTTATCCCGGAGGCTACGG 4064
 Db 104078 TCAACTGGGAAACACTGCATATTCGCGCAGGCGGAGTTATCCCGGAGGCTACGG 104019
 Qy 4065 ACTTGATAAGGAGGCTCTGTGCTGCGCTGACAAAGCGGCTGGGCAAGAGCTGGACGAGG 4124
 Db 104018 ACTTGATAAGGAGGCTCTGTGCTGCGCTGACAAAGCGCTGGGCAAGAGCTGGACGAGG 103959

Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacle, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Swirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissbach, J., Williams, S.M., Woodage, T., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of *Drosophila melanogaster*.
Science 287 (5461), 2185-2195 (2000)
20196006
10731132
2 (bases 1 to 236109)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, J.C.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
FEATURES
Source
Location/Qualifiers
1..236109
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
/contig="CG1342"
/gene="CG1342"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
join(20032..21133,21223..22071)
/gene="CG1342"
/product="CT3018"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
join(20033..21133,21223..22071)
/gene="CG1342"
/note="CG1342 gene product"
/codon_start=1
/protein_id="AAF57075.1"
/db_xref="GI:7301970"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
/translation="MKAVLCATFTLAAIGQGLTQLEDENHSGFAGOVSLIALQL
LKFNKIDANQVHSPGLGVSILAMLAESAEGDYSEFOVGVGPKDRTKLRDYKRIL
GSYQNRDAVALPSPFTWLYIYRNSAREFKDLKHVYDVVDKISRQEDWNEPT
SLQSEKNTDSTEGSTEOSKDVIGETLKRILDDIEDIPAIFSDGQGIIDKE
ASKFDVDQKVQVPAQAEALLOKQKQATTESESOPEETTTLSVEKOEK
DMAAENPAEKQKSDQESQIKNEENETVQEEKLAKIMAAPALTAGPEKVR
PQKLENAVTAADGADEIMALESHLPFSVRNGARSLFQDDITSAISANSITGR
SAGSKMLLFLNGLYRGSWANFYQLRDGSDEFFMTNEDAVKAPMHARGKFOVAD
LPQKARVLLPETSRYALCIVLPDETEGLSDVI SOLQTSDFLAKQFQKELHIS
MPKFQVETSRSEAMLKQGLKKVSRTEAQLSLSEDVHVDVDELQVNVNVRVDEGG
SSANSLSAATMQARTPSVESTPLVPPEPELPVGFVERFVNRPFATFIVDCEQEVLA
SGLIYPERKEDLPSSVIEVEEQS"
complement(<23504..>24751)
/gene="CG12069"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
complement(join(<23504..23717,23771..24048,24117..24282,
24339..>24751))
/gene="CG12069"
/product="CT460"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
complement(join(23504..23717,23771..24048,24117..24282,
24339..24751))
/gene="CG12069"

/note="CG12069 gene product"
/codon_start=1
/protein_id="AAF57076.1"
/db_xref="GI:7301971"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
/translation="MGQPQQAQMHFSPKVDYIILLDKLREDFNKKFATWTPSPSTGL
DDYELKATLGSFGKQVLVRESEGGYIAQSKQSKDOIIVTKQVSHVSEKNNVLRSM
TFPYNLILASYKDFSLYLDVPLIGGELFTYHRKYRKEFEQAQYAAQVFLALEY
LHSCSLYRLDKPENIMMDKNGYLKDFGFAKVTETMTLCGTPYLPPELTIQSKP
YGTSDWAFGLVFEFVAGHSPFSAHNROVMYKICICADYKMPSYFSSARHLVD
HLLOVLSKRFGLNLINGNRDILKEHEWFKDVEWIPLLNQTYNAPVVPNISPEDISNPD
KVSQKPRKAKTMEFEAFAD"
complement(<24899..>26338)
/gene="Pka-C2"
/note="CG12066"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0012066"
/db_xref="FLYBASE:FBgn0000274"
complement(join(<24899..25229,25297..25574,25629..25794,
25875..26104,26162..>26338))
/gene="Pka-C2"
/product="CT4596"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/db_xref="FLYBASE:FBan0012066"
complement(join(24899..25229,25297..25574,25629..25794,
25875..26104,26162..26338))
/gene="Pka-C2"
/note="Pka-C2 gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=1
/protein_id="AAF57077.1"
/db_xref="GI:7301972"
/db_xref="FLYBASE:FBan0012066"
/db_xref="FLYBASE:FBgn0000274"
/translation="MSQHTSOYFNSKEDYNLLDNMSREFERNHQTQSPYTNLEN
YITRAVLGNSFGTYMLVREKSGKYAAKMSKDELVRKQVAVHNEKHVNAARE
PFLIYLVSTKCFDYILPLVNGGELFSYHRRVRKNEKHFAYAAQVALALYMH
KMLHMYRDLKPNILLDQRYIKITDFGTRKVDGRTSLCGTPYLPPELTIQSKP
KSDWMAFGILYEFVAGHSPFSAHNROVMYKICICADYKMPSYFTEQVLRVPSL
MQVDTSKRKLERRLRRESSVPGRLVHWSQPSGSHRPLPAHFRRRSVRLRLRV
QGSQVQPNKPPSRIVCEFLNVNVSFEIVFVSVCLISRGAIIVIGFCKSFRTALLH
S"
complement(<5458..>57119)
/gene="CG1340"
/map="100A5-100A5"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
join(5458..56524,56585..>57119)
/gene="CG1340"
/product="CT3002"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
join(55680..56524,56585..>57119)
/gene="CG1340"
/note="CG1340 gene product"
/codon_start=1
/protein_id="AAF57078.1"
/db_xref="GI:7301973"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
/translation="MANRSQDRREFNRDNTYNNRNRDNRHNRDNRNRENSFNQY
RNRDRDRQROVTEPEFIAYVGNLPGVLGVQGYMKIFSEFVKVNRLLIKDRETFE
KGYGVETEIAQLKLSALNCGRKLKLNFSAPLRIDIAHRRQNPAGPVGVGAPPAG
VCHERGCGVGGGSGTGNRNNPYORNNRYRDRDSDVSGHQRFRREPSNSNHQMSNS
TYNFRMNSNSGGGAGRGHGRGPNFNGGGGPGFVNVNGGNTNFVONNRDR
RGHYNPHSGSGSGYNNNNNTFGGNNHPVHGASASGCTGPPSSNOAVRDDDR
PKLVLPRTVTPINSLAETKQALIFGKAKPRDSDSTPVPSSPQDLDASSGAPSDH
DHDFGLGN"

gene

complement(<59297. >60516)
/gene="CG11313"
/map="100A5-100A5"
/db_xref="FLYBASE:FBan0011313"
/db_xref="FLYBASE:FBgn0039798"
complement(join(<59297. 59885, 59939. 60401, 60456. >60516))
/gene="CG11313"
/product="CT31571"
/db_xref="FLYBASE:FBan0011313"
/db_xref="FLYBASE:FBgn0039798"
complement(join(59297. 59885, 59939. 60401, 60456. 60516))
/gene="CG11313"
/note="CG11313 gene product"
/codon_start=1
/protein_id="AAF57079.1"
/db_xref="GI:7301974"
/db_xref="FLYBASE:FBan0011313"
/db_xref="FLYBASE:FBgn0039798"
/translation="MKVIAVLLCLLIIRTAHQGVVSCRPNQRTGYCVNIPLCVPLN

Query Match 68.9%; Score 3942.2; DB 3; Length 236109;
Best Local Similarity 91.6%; Pred. No. 0;
Matches 4444; Conservative 22; Mismatches 17; Indels 370; Gaps 13;

QY 1223 CAGAAATTACACACCTCTCGGATACACGGGACCAACGAGCAAGGATGCACTTACTCT 1282
DB 123864 CAGAAATTACACACCTCTCGGATACACGGGACCAACGAGCAAGGATGCACTTACTCT 123805
QY 1283 GACTATCACCAGCAAGCAGCGATGGAGCGGCACACCTCGGCTCTCTCTCGTTCGGAC 1342
DB 123804 GACTATCACCAGCAAGCAGCGATGGAGCGGCACACCTCGGCTCTCTCTCGTTCGGAC 123745
QY 1343 GTGGTTCATACCGCGCGCGCCCTTGTAGTTCAGCGCGGAGCGGCTCCATATCCGTA 1402
DB 123744 GTGGTTCATACCGCGCGCGCCCTTGTAGTTCAGCGCGGAGCGGCTCCATATCCGTA 123685
QY 1403 TCCGCTGTGGGCTTGGAGTGTGGTGTGGCGCAAGGAGCTGTSCCAAGATGATGACG 1462
DB 123684 TCCGCTGTGGGCTTGGAGTGTGGTGTGGCGCAAGGAGCTGTSCCAAGATGATGACG 123625
QY 1463 GCCTTAATCCCAACAACTGATCCGAGAGCGGAGCATGCAACGGGACACGGGAGCACT 1522
DB 123624 GCCTTAATCCCAACAACTGATCCGAGAGCGGAGCATGCAACGGGACACGGGAGCACT 123565
QY 1523 CACTACCTCGCTCGAGTCCGGCTCTGGACTCGGAGCGGCTAGCTCCCGATCGGACAGC 1582
DB 123564 CACTACCTCGCTCGAGTCCGGCTCTGGACTCGGAGCGGCTAGCTCCCGATCGGACAGC 123505
QY 1583 CCCCATTGCAACACACACCGGAGCTCGAGGACGGTGGTAAATCCAGGTGGAAT 1642
DB 123504 CCCCATTGCAACACACCGGAGCTCGAGGACGGTGGTAAATCCAGGTGGAAT 123445
QY 1643 GGTGATTTTCCTGCTGCGCAAGCGTTTCAGTGGTGGCTCCACCGGCGCGCGGCA 1702
DB 123444 GGTGATTTTCCTGCTGCGCAAGCGTTTCAGTGGTGGCTCCACCGGCGCGCGGCA 123385
QY 1703 CGCAATCCACCGCTSCAGCGCGCCACCGCCCGCCACCGCCAGTCCCGCCACCGCCAG 1762
DB 123384 CGCAATCCACCGCTSCAGCGCGCCACCGCCCGCCACCGCCAGTCCCGCCACCGCCAG 123325
QY 1763 GGTACGTGAACGGCGATCACCGGCGCTTGAACACCGCGCGCGCGGATAGCGCACCC 1822
DB 123324 GGTACGTGAACGGCGATCACCGGCGCTTGAACACCGCGCGCGGATAGCGCACCC 123265
QY 1823 ACTAGCGGGCACTACCTGTATATACCCARACGGGCTGAAGAACCCCGCAGCAGCAG 1882
DB 123264 ACTAGCGGGCACTACCTGTATATACCCARACGGGCTGAAGAACCCCGCAGCAGCAG 123205
QY 1883 TTGACGACGACCTGAAGTCCCTGAACCTATACCCAGCGGAGGAGTGGACGATGTG 1942
DB 123204 TTGACGACGACCTGAAGTCCCTGAACCTATACCCAGCGGAGGAGTGGACGATGTG 123145
QY 1943 GAGCCACCGCGCCCTACCTAATTCAGGCGGAGCGGAGGAGCAGCAGCGCGCGGCA 2002

DB 123144 GAGCCACCGCGCCCTACCTAATTCAGGCGGAGCGGAGGAGCAGCAGCGCGCGCA 123085
QY 2003 CCACCCAGTTACAGCGCTTCATGAGTCGCGGCGAGTCCGCCACACAAATCCCAAAATCG 2062
DB 123084 CCACCCAGTTACAGCGCTTCATGAGTCGCGGCGAGTCCGCCACACAAATCCCAAAATCG 123025
QY 2063 GACTACAGAAATCCCGAGCAGTGGGATATATCTCGGCCACCTCGGCGGCTCGCGGAGC 2122
DB 123024 GACTACAGAAATCCCGAGCAGTGGGATATATCTCGGCCACCTCGGCGGCTCGCGGAGC 122965
QY 2123 CCATAACTGTCTGCTGCGCGCGCGCTGGGGAAGCCACAAACAGAGTCTTACACAG 2182
DB 122964 CCATAACTGTCTGCTGCGCGCGCGCTGGGGAAGCCACAAACAGAGTCTTACACAG 122905
QY 2183 GCCAGAGTCAGCAGCGGATCATGATGAGAGTGTGAAGAGCAGCAGGTCCAAAGGCC 2242
DB 122904 GCCAGAGTCAGCAGCGGATCATGATGAGAGTGTGAAGAGCAGCAGGTCCAAAGGCC 122845
QY 2243 GTGCTGCAACAGCAGTGGCGGCCCAATCGCATCGAGTGCCTCGGCCAGCAATTCACCA 2302
DB 122844 GTGCTGCAACAGCAGTGGCGGCCCAATCGCATCGAGTGCCTCGGCCAGCAATTCACCA 122785
QY 2303 GTCCAGTGTCTGCGCGCTCCACCTCTTACCCTCAGAAAGTCCGCGGAGTGTGACAG 2362
DB 122784 GTCCAGTGTCTGCGCGCTCCACCTCTTACCCTCAGAAAGTCCGCGGAGTGTGACAG 122725
QY 2363 CAGCAACAGGAGCAGCGCGGCCACACAGCAGCAGCAGCAGCAGCAATTCACCA 2422
DB 122724 CAGCAACAGGAGCAGCGCGGCCACACAGCAGCAGCAGCAGCAGCAATTCACCA 122665
QY 2423 CCAAGCGCCMACCACACCGCCCTTGGTGGTCTGAAYAGCAAGCCCAATTCCTGGAGCA 2482
DB 122664 CCAAGCGCCACACACCGCCCTTGGTGGTCTGAAYAGCAAGCCCAATTCCTGGAGCA 122605
QY 2483 CCGTCTTATGCAAGAGTATGAGCCAAAGCGGCGCACGGTGGTATGAGCAGCAGCAA --- 2539
DB 122604 CCGTCTTATGCAAGAGTATGAGCCAAAGCGGCGCACGGTGGTATGAGCAGCAGCAA 122545
QY 2540 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2599
DB 122544 CAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 122485
QY 2600 CAACGAACTGCAGCGCTTGGAGGTCCTCAGGCGCACAGGCTCAGA - - - - - 2645
DB 122484 CAACGAACTGCAGCGCTTGGAGGTCCTCAGGCGCACAGGCTCAGA - - - - - 122425
QY 2646 -GGGAGCGGATCAACGGGAGCGGGAACGGGATCAGCAGAACTGGCCAAACGGAATCT 2704
DB 122424 CCGGAGCGGATCAACGGGAGCGGGAACGGGATCAGCAGAACTGGCCAAACGGAATCT 122365
QY 2705 GCGCGCAGATGCTTCCCGCGCGCGCTTATCAGGCAACAAACAGCAGCAGATC 2764
DB 122364 GCGCGCAGATGCTTCCCGCGCGCGCTTATCAGGCAACAAACAGCAGCAGATC 122305
QY 2765 AAACCGCGGAGTGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 2824
DB 122304 AAACCGCGGAGTGAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 122245
QY 2825 CCACCCATTCGCGCTGCCAAATACAAATACAACTCTCCAAACAGCGGCGCAATAGCTCG 2884
DB 122244 CCACCCATTCGCGCTGCCAAATACAAATACAACTCTCCAAACAGCGGCGCAATAGCTCG 122185
QY 2885 GCGCGGAGCAAGGATTCACCGGCGCGCGCTTCTCTGTCAGCAGTTCGCAAGAGATC 2944
DB 122184 GCGCGGAGCAAGGATTCACCGGCGCGCGCTTCTCTGTCAGCAGTTCGCAAGAGATC 122125
QY 2945 AAGCAGCGCTCGCCCTCCGAGGCGCAAGAGATCTCCAAAGGAGAGGAGGAGCGC 3004
DB 122124 AAGCAGCGCTCGCCCTCCGAGGCGCAAGAGATCTCCAAAGGAGAGGAGGAGCGC 122065
QY 3005 AAGGAGTTCGCGATCAGGCGAGTACTCGCGCGCAAGCTTCAAGTTCCTTATGAGGAGCAG 3064

Db 122064 AAGGAGTTCCGCATCAGGCGAGTACTCGCCGCAAGCCTTCAAGTTCTTCATGCGAGCAGCAC 122005
QY ATAGAGAACCTGATCAAGTCTGATATCGCCAGCGCACGATATCGAAGAAATCAGCTGGAGAG 3124
Db 122004 ATAGAGAACCTGATCAAGTCTGATATCGCCAGCGCACGATATCGAAGAAATCAGCTGGAGAG 121945
QY 3125 GAGATGCACAAAGTGGACTGCGCGATCAGACCCAAATCGAGATGAGGAAAAATGCTGAAC 3184
Db 121944 GAGATGCACAAAGTGGACTGCGCGATCAGACCCAAATCGAGATGAGGAAAAATGCTGAAC 121885
QY 3185 CAAAGGAGAGCAACTACATTCGATTGAAGCGCGCCAAAGATGAGCAAGAGCATGTTCTGTC 3244
Db 121884 CAAAGGAGAGCAACTACATTCGATTGAAGCGCGCCAAAGATGAGCAAGAGCATGTTCTGTC 121825
QY 3245 AAATGAAAGCCANTGGAGTGGGTGCATTTGGCGAGGTAACGCTGGTGAGCAAAATCGAT 3304
Db 121824 AAATGAAAGCCANTGGAGTGGGTGCATTTGGCGAGGTAACGCTGGTGAGCAAAATCGAT 121765
QY 3305 ACCTCGAACCAATTTGATGCGATGAAAACCCCTCGGGAAGCGGAGCTTCTCAAGCGCAAT 3364
Db 121764 ACCTCGAACCAATTTGATGCGATGAAAACCCCTCGGGAAGCGGAGCTTCTCAAGCGCAAT 121705
QY 3365 CAGGTGGCACACGTGAAGCGCGAGAGGATATCCTCGCGGAAGCGCAATAAATCGGTG 3424
Db 121704 CAGGTGGCACACGTGAAGCGCGAGAGGATATCCTCGCGGAAGCGCAATAAATCGGTG 121645
QY 3425 GTGAAGTTGTAATACAGCTTCCAGGACAAAGATAATCTGTACTTTGTCATGGACTACATA 3484
Db 121644 GTGAAGTTGTAATACAGCTTCCAGGACAAAGATAATCTGTACTTTGTCATGGACTACATA 121585
QY 3485 C----- 3485
Db 121584 CCAGGTGAGTGGAAAACCTTAAGCAGTCCAAAGATGTGATCGGAGTGCATAATGCTCATTTG 121525
QY 3486 -----CAGGTGGTGTATCTGATCTGCTCTCATCAAACTGGGCATTTTCGAGGAGG 3536
Db 121524 GATTGTCGACAGGTGGTGATCTGATCTGCTCATCAAACTGGGCATTTTCGAGGAGG 121465
QY 3537 AACTGGCCGAGTTCTACATCGCCGAGGTCACTGCGCGGTGGACAGCGTTTCACAAAATGG 3596
Db 121464 AACTGGCCGAGTTCTACATCGCCGAGGTCACTGCGCGGTGGACAGCGTTTCACAAAATGG 121405
QY 3597 GTTTCATT----- 3604
Db 121404 GTTTCATTACAGGTAATTAGCAGTTGCTTCAGTTATACAGTCAATTTACATATTGT 121345
QY 3605 -----CACAGAGATCAAGCTGACAACTACTCATCGATAGGAGCGGACACA 3653
Db 121344 ATGGATTATCCACAGAGACATCAAGCTGACAACTACTCATCGATAGGAGCGGACACA 121285
QY 3654 TAAAGCTCACCGACTTTGGCTGTGCACGGGATTCGATGGAGCGCAACTCGAAGTACT 3713
Db 121284 TAAAGCTCACCGACTTTGGCTGTGCACGGGATTCGATGGAGCGCAACTCGAAGTACT 121225
QY 3714 ACCAGGAGAAC----- 3724
Db 121224 ACCAGGAGAAGCGTGATGGAACCCCGACTAACGTAGTGTAAACAGCTTTAATAATAT 121165
QY 3725 -----GGCAATCACTCGCGCAGACTCGATGAGCCCTGGGAGGAATA 3768
Db 121164 CATTCTACTCTGTATAGCAATCACTCGCGCAGACTTCGATGAGCGCCTGGGAGGAATA 121105
QY 3769 CTCGGAGAACGACCGAAGCCCAACCGCTGGAG----- 3802
Db 121104 CTCGGAGAACGACCGAAGCCCAACCGCTGGAGAGGTTGAGTTTGGTACACTTAATTAAT 121045
QY 3803 -----AGCGGAGGATGCGCATCACCACA 3827
Db 121044 AGGATCGGCGGTTAATAAAAGTATGCGGAACCTCGCAGCGCGGATGCGCATCACCACA 120985
QY 3828 GAGTCTGGCCCACTCGCTGGTGGCACCCCGAACTACATAGCTCCCGAGGTGCTGGAGA 3887
Db 120984 GAGTCTGGCCCACTCGCTGGTGGCACCCCGAACTACATAGCTCCCGAGGTGCTGGAGA 120925

QY	3888	GRATGGGTACAGCAGCTGTCGACTTACTGGAGCGTGCGCGTCACTCTTTAYGAGATCC	3947
Db	120924	GAACTGGGTACACGCAGCTGTGGCACTACTGGAGCGTGCGCGTCACTCTATGAGATCC	120865
QY	3948	TGGTGGGTACAGCAGCCCTTTCTTGCCCAACAGTCGCTGGAAACGCCAACAAA-----	3998
Db	120864	TGGTGGGTACAGCAGCCCTTTCTTGCCCAACAGTCGCTGGAAACGCCAACAAAAGTTCGCT	120805
QY	3999	-----AGGTCA	4004
Db	120804	CTTTGGATTCTCTATAAAGCCAAGTCATTACACAATGACTTCTTCAITTAACAGGTCA.	120745
QY	4005	TCAACTGGGAGAAACSCSTGCTATATTCGCGCAGCGCGAGTATCCCGCAGGCTACGG	4064
Db	120744	TCAACTGGGAGAAACCCTGCTATATTCGCGCAGCGCGAGTATTCGCGCAGGCTACGG	120685
QY	4065	ACTTGATAAGGAGGCTCTGTGGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACGAGG	4124
Db	120684	ACTTGATAAGGAGGCTCTGTGGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACGAGG	120625
QY	4125	TCAAGAGCCACGACTTCTTCAAGGCGATCGACTTTGCGGACATCGGNAGCAGAAACGC	4184
Db	120624	TCAAGAGCCACGACTTCTTCAAGGCGATCGACTTTGCGGACATCGGNAGCAGAAACGC	120565
QY	4185	CCTACATACCGGAATCAAGCACCCRACGACACATCCAACCTTTGATCCCGTGGATCCGG	4244
Db	120564	CCTATATACCGGAATCAAGCACCCRACGACACATCCAACCTTTGATCCCGTGGATCCGG	120505
QY	4245	AGAAGCTGGCTCCGAATGACTCCACCATGAGCAGCGGCGATGATGTCCACGANTGACC	4304
Db	120504	AGAAGCTGGCTCGAATGACTCCACCATGAGCAGCGGCGATGATGTCCACGANTGACC	120445
QY	4305	GCACYTTCCACGGCTTTTTCGAATTTACTTCGCTCGCTTCTTCGACGACAAAGCGCCG	4364
Db	120444	GCACYTTCCACGGCTTTTTCGAATTTACTTCGCTCGCTTCTTCGACGACAAAGCGCCG	120385
QY	4365	CGGATATGACGGAGCATCAGCGCGCGTTTACGCTCTCAAATGGATGCTCTCATGTGCC	4424
Db	120384	CGGATATGACGGAGCATCAGCGCGCGTTTACGCTCTCAAATGGATGCTCTCATGTGCC	120325
QY	4425	AACACCAACACCCCGCCCGGATCATTTAGTCAAAATAGTCACAAAAAGGGGATAGAA	4484
Db	120324	AACACCAACACCCCGCCCGGATCATTTAGTCAAAATAGTCACAAAAAGGGGATAGAA	120265
QY	4485	ACCATTGAGTGGGCTGCATTGTAAAGGAAGCGTGGCTATAGAAATGAACTACTATATA	4544
Db	120264	ACCATTGAGTGGGCTGCATTGTAAAGGAAGCGTGGCTATAGAAATGAACTACTATATA	120205
QY	4545	CATTATATAATTAGAGACACTAGAGCGGGAGCTAGCTATATATACATACAAAATAATA	4604
Db	120204	CATTATATAATTAGAGACACTAGAGCGGGAGCTAGCTATATATACATACAAAATAATA	120145
QY	4605	TACATATATTTGATATATATATATATATATATGCG- GTAGGGCATGAATGAATAATAT	4663
Db	120144	TACATATATTTGATATATATATATATATATATATATATGCG- GTAGGGCATGAATGAATAATAT	120085
QY	4664	AAAACGAGCGCGAGTAGAGTGAAG- CGAGAGGAGCGAGTCAGGACCTTCGACCTTTAAC	4722
Db	120084	AAAACGAGCGCGAGTAGAGTGAAGACCGAGAGGAGCGAGTCAGGACCTTCGACCTTTAAC	120025
QY	4723	TGACATAGTATATCTTGTGCACACTACTCTCCACAAATAATATATTTTTTAATTTGTT	4782
Db	120024	TGACATAGTATATCTTGTGCACACTACTCTCCACAAATAATATATTTTTTAATTTGTT	119965
QY	4783	AGAAATCAAAGGGRCCAACTGGAAATCGAACTTTCTGTGCTCTCAAGCAAGCAAGC	4842
Db	119964	AGAAATCAAAGGGRCCAACTGGAAATCGAACTTTCTGTGCTCTCAAGCAAGCAAGC	119915
QY	4843	AAAGCAAAACAAAACGCTTTAAACTAAAYGAGACGCGAAATTTACCCAACCACTTCAC	4902
Db	119914	AAAGCAAAACAAAACGCTTTAAACTAAAGCAGCGGAATTTACCCCAACCACTTCAC	119855

QY 4903 TCTCCTTTCTCCACCTCCGATCGTGGCCGGATTCGAACCTCAGAGCTGGTTGCATCG 4962
 Db 119854 TCTCCTTTCTCCACCTCCGATCGTGGCCGGATTCGAACCTCAGAGCTGGTTGCATCG 119795
 QY 4963 GCCATCCCATTTACCTCCATTCAGAAATTCAGATTCGAGGTGGTGGAGAGCAAC 5022
 Db 119794 GCCATCCCATTTACCTCCATTCAGAAATTCAGATTCGAGGTGGTGGAGAGCAAC 119735
 QY 5023 GGAGCAACAAAGTCGCGACGCGAGCATATAAGCGGGTCTTTAAGCCTTAATCTAAATCTA 5082
 Db 119734 GGAGCAACAAAGTCGCGACGCGAGCATATAAGCGGGTCTTTAAGCCTTAATCTAAATCTA 119675
 QY 5083 AACTGGGAGAACAGGACCCATGTGTAATAGTGTCTCTATCCCAATTCGCTATC 5131
 Db 119674 AACTGGGAGAACAGGACCCATGTGTAATAGTGTCTCTATCCCAATTCGCTATC 119615
 QY 5132 ACTGCTCTTCANCTGWTAGGACCCCGCCCGCCCTC-----CCCATCCAAAGAACAA 5187
 Db 119614 ACTGCTCTTCANCTGWTAGGACCCCGCCCGCCCTC-----CCCATCCAAAGAACAA 119555
 QY 5188 ACTTAGACCTAGCTATGTGAAAGCTAGCAATGTTAGCAACTTGTGTAATGCCAAAT 5247
 Db 119554 ACTTAGACCTAGCTATGTGAAAGCTAGCAATGTTAGCAACTTGTGTAATGCCAAAT 119495
 QY 5248 GAAATGTTTACCCCGACGAGGAAACGCGGGGAAATTTCAACACTTATCTCTGATAGC 5307
 Db 119494 GAAATGTTTACCCCGACGAGGAAACGCGGGGAAATTTCAACACTTATCTCTGATAGC 119435
 QY 5308 AAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5367
 Db 119434 AAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 119380
 QY 5368 TTAATGTAATATTGTAAGAACACGTTAATGTAATGTAATGTAATGTAATGTAATG 5427
 Db 119379 TTAATGTAATATTGTAAGAACACGTTAATGTAATGTAATGTAATGTAATGTAATG 119320
 QY 5428 TAAGATGTTTATGTTATAGACCGCTAACCGTAACTTAGTAACTTAACTTAACTTAACT 5487
 Db 119319 TAAGATGTTTATGTTATAGACCGCTAACCGTAACTTAGTAACTTAACTTAACT 119260
 QY 5488 CGAGAGTACAGTACATGTTGTTTGTGTCGAGTGTGTCGAGTGTGTCGAGTGTGTCG 5547
 Db 119259 CGAGAGTACAGTACATGTTGTTTGTGTCGAGTGTGTCGAGTGTGTCGAGTGTGTCG 119200
 QY 5548 GAAACGATTTGTTTCTCTTTAAATAGCTTCAGTTGTTGATGCTGCTGCTGCTGCTG 5607
 Db 119199 GAAACGATTTGTTTCTCTTTAAATAGCTTCAGTTGTTGATGCTGCTGCTGCTGCTG 119140
 QY 5608 GACTTATATATAGTCCATCTGAATATTCGTTGATGGAGCTATTTTAAATGTGAGATCGA 5667
 Db 119139 GACTTATATATAGTCCATCTGAATATTCGTTGATGGAGCTATTTTAAATGTGAGATCGA 119080
 QY 5668 GCTAATGGAAGAAATACAAACAACTCTGTCGCTGGCCCAATTAGTTTAC 5720
 Db 119079 GCTAATGGAAGAAATACAAACAACTCTGTCGTGCTGGCCCAATTAGTTTAC 119027

RESULT 6
 AC012975/c 28133-bp
 LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 ACCESSION AC012975
 VERSION AC012975.1
 KEYWORDS HTG; HGFS_PHASE2
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila
 REFERENCE 1 (bases 1 to 28133)
 AUTHORS Adams,M. and Venter,J.C.
 TITLE Direct Submission

JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
 COMMENT This sequence was identified as CDM:10211994 by the submitter. For further information on this sequence you may e-mail to: fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This finished sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source location/Qualifiers
 1..28133
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 7406 a 6301 c 6224 g 8202 t
 ORIGIN
 Query Match 51.8%; Score 2965.2; DB 2; Length 28133;
 Best Local Similarity 90.3%; Pred. No. 0;
 Matches 3369; Conservative 16; Mismatches 9; Indels 338; Gaps 7;
 QY 1223 CAGAAATTACACACCTCTGCGATACAGCGGACCAAGGACGCAAGCATGCTACTCT 1282
 Db 3732 CAGAAATTACACACCTCTGCGATACAGCGGACCAAGGACGCAAGCATGCTACTCT 3673
 QY 1283 GACTATCACACGCGCAAGCAGCGGATGAGCGCGCCACCTCCGCTCTCTCTCGGAC 1342
 Db 3672 GACTATCACACGCGCAAGCAGCGGATGAGCGCGCCACCTCCGCTCTCTCTCGGAC 3613
 QY 1343 GTGGTCATACCGCGCGCGCCGCTTTAGTGTACGCGCGGCGGCTTCCATATCCGTA 1402
 Db 3612 GTGGTCATACCGCGCGCGCGCCGCTTTAGTGTACGCGCGGCGGCTTCCATATCCGTA 3553
 QY 1403 TCCGCTGTGGCGGTGGAGTGTGGTGTGGCGAAGCAGGAGTGTSCCAAGATGATGAGC 1462
 Db 3552 TCCGCTGTGGCGGTGGAGTGTGGTGTGGCGAAGCAGGAGTGTSCCAAGATGATGAGC 3493
 QY 1463 GCCTTAATCCCAACAACTGATCGGAAGCGAGCATCGAAGCGGACGACGCGAGCAGT 1522
 Db 3492 GCCTTAATCCCAACAACTGATCGGAAGCGAGCATCGAAGCGGACGACGCGAGCAGT 3433
 QY 1523 CACTACCTGCGTGCAGTCCGGCTCTGGACTCCGGAGCGGTAGTCCCGATCGGACAGC 1582
 Db 3432 CACTACCTGCGTGCAGTCCGGCTCTGGACTCCGGAGCGGTAGTCCCGATCGGACAGC 3373
 QY 1583 CCGCATTCGCACACACACCCAGCCAGCTCGAGGACGCTGGTAAATCCAGGTGGAAT 1642
 Db 3372 CCGCATTCGCACACACACCCAGCCAGCTCGAGGACGCTGGTAAATCCAGGTGGAAT 3313
 QY 1643 GGTGATTTTCTCCGTCGCAAGCGGTTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 1702
 Db 3312 GGTGATTTTCTCCGTCGCAAGCGGTTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 3253
 QY 1703 CGCAATCCACCGCTTSCAGCGCGCCAGCCGCCCCACCGCCAGTGCCTGCGCCACAGCCAG 1762
 Db 3252 CGCAATCCACCGCTTSCAGCGCGCCAGCCGCCCCACCGCCAGTGCCTGCGCCACAGCCAG 3193
 QY 1763 GGTGATTCGAGCGGATCACCGGCTTGAACACCGCGCGCGCGCGATAGCGCCACCC 1822
 Db 3192 GGTGATTCGAGCGGATCACCGGCTTGAACACCGCGCGCGCGATAGCGCCACCC 3133
 QY 1823 ACTCAGCGGCGCAACTCACCTGTAATACCCAAACGGGCTCAAGAACCCGCGAGCAGAG 1882
 Db 3132 ACTCAGCGGCGCAACTCACCTGTAATACCCAAACGGGCTCAAGAACCCGCGAGCAGAG 3073
 QY 1883 TTGACGCGAGCAGTGAAGTCCCTGAACCTTATACCCAGCGGAGGAGTGGAGCTGTG 1942
 Db 3072 TTGACGCGAGCAGTGAAGTCCCTGAACCTTATACCCAGCGGAGGAGTGGAGCTGTG 3013
 QY 1943 GAGCCACCGCGCGCTACCTAATTCAGCGGAGCGCGGAGGAGCAGCCGCGCGCGCA 2002
 Db 3012 GAGCCACCGCGCGCTACCTAATTCAGCGGAGCGCGGAGGAGCAGCCGCGCGCGCA 2953
 QY 2003 CCACCCAGTTACACGGGCTTCCATGTCAGTGCAGTGCAGCCACACAAATCCCAACATCG 2062

Db 1872 ATAGAAACGTGATCAAGTCGTATGCCAGCGCACCTATCGCAAGAAATCAGCTCGAGAAG 1813

QY 3125 GAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATGCTGAAC 3184

Db 1812 GAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATGCTGAAC 1753

QY 3185 CAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGACAAGAGCATGTTGCGTC 3244

Db 1752 CAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCCAGATGGACAGACATGTTGCGTC 1693

QY 3245 AAACCTGAAGCCCATGGAGTGGGTGCATTTGGCGAGGTAAACGCTGGTGAGCAAAATCGAT 3304

Db 1692 AAACCTGAAGCCCATGGAGTGGGTGCATTTGGCGAGGTAAACGCTGGTGAGCAAAATCGAT 1633

QY 3305 ACCTCGAACCACTTTGTATGCCATGAAACCCCTCGGAAAGCGGACGCTTCACAGCCGAAT 3364

Db 1632 ACCTCGAACCACTTTGTATGCCATGAAACCCCTCGGAAAGCGGACGCTTCACAGCCGAAT 1573

QY 3365 CAGGTGGCACAGTGAAGGCGGAGAGGATATCCTCGCGAAGCCGACAAATACTGGGTG 3424

Db 1572 CAGGTGGCACAGTGAAGGCGGAGAGGATATCCTCGCGAAGCCGACAAATACTGGGTG 1513

QY 3425 GTGAAGTGTACTACACTTCCAGGACAAGGATAAATCTGTACTTTGTGAGTGAACATA 3484

Db 1512 GTGAAGTGTACTACACTTCCAGGACAAGGATAAATCTGTACTTTGTGAGTGAACATA 1453

QY 3485 C----- 3485

Db 1452 CAGGTGAGTGGAAACTTAAGCAGTCAACGATGTGTATCGGAGTCCAATGTCATTTGTG 1393

QY 3486 -----CAGGTGGTATCTCATGTGCTGTCTATCAAACTGGGCATTTTCGAGGAG 3536

Db 1392 GATTGTCGACAGGTGGTATCTCATGTGCTGTCTATCAAACTGGGCATTTTCGAGGAGG 1333

QY 3537 AACTGGCCAGATTCTACATCGCCGAGGTCACTCGCCGTGGACAGCGTTTCACAAAATGG 3596

Db 1332 AACTGGCCAGATTCTACATCGCCGAGGTCACTCGCCGTGGACAGCGTTTCACAAAATGG 1273

QY 3597 GCATCAT----- 3604

Db 1272 GCATCATCACAGTAAATAGCAGTTGCTTCAGTTATCACAGTCACATTTACATATTGTT 1213

QY 3605 -----CACAGAGACATCAAGCCCTGACAACATCTCATCATAGTAGGACGACACACA 3653

Db 1212 ATGGATTATCCACAGACATCAGCCGTGACACATCTCATGTATAGGACGACACACA 1153

QY 3654 TAAAGTCAACGACTTTGGCCTGTGCACGGGATCCGATGGACGACACAACCTCGAAGTACT 3713

Db 1152 TAAAGTCAACGACTTTGGCCTGTGCACGGGATCCGATGGACGACACAACCTCGAAGTACT 1093

QY 3714 ACCAGGAGAAC----- 3724

Db 1092 ACCAGGAGAACGGTGAGATGGNACCCAGCTAAAGTGTAGTGTAAACAGCTTTAATAATAT 1033

QY 3725 -----GGCAATCACTCGCCGACGACTCGATGTAGCGCCCTGGGAGGAATA 3768

Db 1032 CATTCCTACTCTGATAGGCAATCACTCGCCGACGACTCGATGGAGCCCTGGGAGGAATA 973

QY 3769 CTCCCAGACGACCGAAGCCACCGTGTGGAG----- 3802

Db 972 CTCCCAGACGACCGAAGCCACCGTGTGGAGAGTGTAGTTTCGTACACTTAATTAAT 913

QY 3803 -----AGCGACCGATCGCGATCACAAA 3827

Db 912 AGGATCGGCCGTTAATAAAGTATGCCGAATCGAGGCGCGGATCGCGATCACAAA 853

QY 3828 GAGTCTGGCCCACTCGTGTGGGACCCCGAACATACATAGTCCCGAGGTGCTCGAGA 3887

Db 852 GAGTCTGGCCCACTCGTGTGGGACCCCGAACATACATAGTCCCGAGGTGCTCGAGA 793

QY 3888 GAGTGGGTACACGCACTGTGGCACTACTGTGAGCGTGGCGCTATCTCTTAYGAGATGC 3947

Db 792 GAGTGGGTACACGCACTGTGGCACTACTGTGAGCGTGGCGCTATCTCTTAYGAGATGC 733

QY 3948 TGTGGTTCAGCGCCCTTCTTGGCCACAGTCCGCTGGAACGCAACAA----- 3998
 Db 732 TGTGGTTCAGCGCCCTTCTTGGCCACAGTCCGCTGGAACGCAACAAAGGTTTCGCT 673
 QY 3999 -----AGTCA 4004
 Db 672 CTTTGGATTCTTATTAAGCAAGTGCATTAACAATGACTTGTCTTATACAGTCA 613
 QY 4005 TCAACTGGGAGAAACSCGTGATATTCGCGCAGCGCCGAGTTATCCGCGAGGCTACGC 4064
 Db 612 TCAACTGGGAGAAACSCGTGATATTCGCGCAGCGCCGAGTTATCCGCGAGGCTACGC 553
 QY 4065 ACTTGATAAGGAGGCTCTGTGGCTGGCTGACAAGCGGTGGGCAAGCGTGGACGAGG 4124
 Db 552 ACTTGATAAGGAGGCTCTGTGGCTGGCTGACAAGCGGTGGGCAAGCGTGGACGAGG 493
 QY 4125 TCAAGAGCCAGGACTTCTTCAAGGGGATCGACTTTCGGGACATGCGGGAAGCGCAAGCGC 4184
 Db 492 TCAAGAGCCAGGACTTCTTCAAGGGGATCGACTTTCGGGACATGCGGGAAGCGCAAGCGC 433
 QY 4185 CTTACATACCGGAAATCAAGCAGCCACACATCCCACTTTGATCCCGTGGATCCGG 4244
 Db 432 CTTATATACCGGAAATCAAGCAGCCACACATCCCACTTTGATCCCGTGGATCCGG 373
 QY 4245 AGAAGCTGGCTCGAATGACTCCACATGAGCAGCGCGGATGATGTCGACCAAGATGACC 4304
 Db 372 AGAAGCTGGCTCGAATGACTCCACATGAGCAGCGCGGATGATGTCGACCAAGATGACC 313
 QY 4305 GCACATTCACCGGCTTTTCGAATTTACCTTCGCTGCTTTCGAGCAGCAGCAGCGC 4364
 Db 312 GCACATTCACCGGCTTTTCGAATTTACCTTCGCTGCTTTCGAGCAGCAGCAGCGC 253
 QY 4365 CGGATATGAGCAGGATCAGCGCGGTTTACGCTGAAATGATGATGCTCTCCATGTCGC 4424
 Db 252 CGGATATGAGCAGGATCAGCGCGGTTTACGCTGAAATGATGATGCTCTCCATGTCGC 193
 QY 4425 AACACCAACACCCCGCCCGAATCATTTAGTCAAAATAGTCAACAAAGGGGATAGAA 4484
 Db 192 AACACCAACACCCCGCCCGAATCATTTAGTCAAAATAGTCAACAAAGGGGATAGAA 133
 QY 4485 ACCATTGAGTGGCTTGCAATTTAAAGGAGCGTGTGCTATAGAACTATCTATATA 4544
 Db 132 ACCATTGAGTGGCTTGCAATTTAAAGGAGCGTGTGCTATAGAACTATCTATATA 73
 QY 4545 CATTATATAATATAGGAGCAGTAGAGCGGAGCTACGTATATACATAAATAATA 4604
 Db 72 CATTATATAATATAGGAGCAGTAGAGCGGAGCTACGTATATACATAAATAATA 13
 QY 4605 TACATATATTG 4616
 Db 12 TACATATATTG 1

RESULT 7
 AC013000/c
 Locus
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 AC013000 pieces
 AC013000.1 GI:6223329
 VERSION HTG: HTGS-PHASE2
 KEYWORDS Drosophila melanogaster
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 39839)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
 COMMENT This sequence was identified as CDM:10212234 by the submitter.

For further information on this sequence you may e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 Location/Qualifiers
 1. 39839
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 11426 a 9282 c 8335 g 10796 t
 ORIGIN
 Query Match 17.2%; Score 981.2; DB 2; Length 39839;
 Best Local Similarity 95.9%; Pred. No. 1.7e-182;
 Matches 1053; Conservative 6; Mismatches 9; Indels 30; Gaps 4;
 QY 4638 CGGTAGGCGATGAACCTGAATAATAATAAAGCGAGCGGAGTAGATGAAGCGAGAGGAG 4697
 Db 39838 CTGTAGGCGATGAACCTGAATAATAATAAAGCGAGCGGAGTAGATGAAGCGAGAGGAG 39779
 QY 4698 CGAGTCAGGACCTTCGACCTTTAACTGAACATAGTATATCTTGTGCACCTACTCTCCAC 4757
 Db 39778 CGAGTCAGGACCTTCGACCTTTAACTGAACATAGTATATCTTGTGCACCTACTCTCCAC 39719
 QY 4758 AACAAATATATATTTTAAATTTAGAAATTCAAAAGGCGCAACTGGAATTCGAACCTT 4817
 Db 39718 AACAAATATATATTTTAAATTTAGAAATTCAAAAGGCGCAACTGGAATTCGAACCTT 39659
 QY 4818 TCTGTGTCCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4877
 Db 39658 TCTGTGTCCT-----CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 39609
 QY 4878 CGAATTTACCAACCACTTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4937
 Db 39608 CGAATTTACCAACCACTTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 39549
 QY 4938 GAACTCAGCAGCGTGGTTCATCCGCGCATCCCATTCCTCCATTCAGAAATTCAGAT 4997
 Db 39548 GAACTCAGCAGCGTGGTTCATCCGCGCATCCCATTCCTCCATTCAGAAATTCAGAT 39489
 QY 4998 GCGAGTGTGCGATGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5057
 Db 39488 GCGAGTGTGCGATGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 39429
 QY 5058 GTCTTATAGCCTTAATCTAAATCTAACTGGGAGACAGGACCTATGTA----- 5106
 Db 39428 GTCTTATAGCCTTAATCTAAATCTAACTGGGAGACAGGACCTATGTAATGTAATAG 39369
 QY 5107 TGTCTCTGCTATCCAAATTCGCTCTATCACTGCTCTTCAWCTGWTGACGCCACCCGCCCC 5166
 Db 39368 TGTCTCTGCTATCCAAATTCGCTCTATCACTGCTCTTCACTGAGTGTGCCCCCTCCCTCC 39309
 QY 5167 C-----CTCCCCATCCAAAAGCAAACTTAGACGTAGCCTATGTGAAAAGCTAGCAATGT 5222
 Db 39308 CTCTCATCCATCCAAAAGCAAACTTAGACGTAGCCTATGTGAAAAGCTAGCAATGT 39249
 QY 5223 TAGACCAACTCTGTGAATCCCAAAATGAAATTTGTAGCCCGCAGGAGAAACCGCGGGA 5282
 Db 39248 TAGACCAACTCTGTGAATCCCAAAATGAAATTTGTAGCCCGCAGGAGAAACCGCGGGA 39189
 QY 5283 AATTCAACACTTATCTCTGATAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAG 5342
 Db 39188 AATTCAACACTTATCTCTGATAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAGCAAAAG 39134
 QY 5343 ACAGTACGAGAGAAATGTAAATCTTCTTAATTAATTTGTAAGAACACAGTTAATGTAA 5402
 Db 39133 ACAGTACGAGAGAAATGTAAATCTTCTTAATTAATTTGTAAGAACACAGTTAATGTAA 39074
 QY 5403 TCTATCTAGACTGTGTAGCGCCCTTAAGATCTTTTATAGCCGCTAACCGTA 5462
 Db 39073 TCTATCTAGACTGTGTAGCGCCCTTAAGATCTTTTATAGCCGCTAACCGTA 39014

QY 5463 ATCTAGTTTAATTCCTAACACTAAACGAGACGTACAGTACATGGTGTGTTTCTGTTGTCGT 5522
DB 39013 ATCTAGTTTAATTCCTAACACTAAACGAGACGTACAGTACATGGTGTGTTTCTGTTGTCGT 38954
QY 5523 AGGTTCTGTTGAAATGCTTAAACGGCAACAGATTTGTTTCTCTTTAAATTAGCTTCAGT 5582
DB 38953 AGGTTCTGTTGAAATGCTTAAACGGCAACAGATTTGTTTCTCTTTAAATTAGCTTCAGT 38894
QY 5583 TTGTATGTCGCTGTTGTTTAAATATGACTTATATATAGTATCCATCTGAATATTCGTTGATG 5642
DB 38893 TTGTATGTCGCTGTTGTTTAAATATGACTTATATATAGTATCCATCTGAATATTCGTTGATG 38834
QY 5643 GAGCCTATTTAAATGTCGAGTCGAGCTAAATGTAAGGAAATACAAACAACTCTGTGTGC 5702
DB 38833 GAGCCTATTTAAATGTCGAGTCGAGCTAAATGTAAGGAAATACAAACAACTCTGTGTGC 38774
QY 5703 CTTGGCCCAATTAGTTTAC 5720
DB 38773 CTTGGCCCAATTAGTTTAC 38756

RESULT 8
LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6359193.
ACCESSION AR201457
VERSION AR201457.1 GI:20252345
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3155)
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
FEATURES
source 1. 3155
BASE COUNT 751 a 924 c. 894 g 586 t
ORIGIN

Query Match 9.98; Score 567.8; DB 6; Length 3155;
Best Local Similarity 64.38; Pred. No.: 3.3e-101;
Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCGGAATAGCTCGGGCGGCGAGCAACGGATCCACCGGACACCGCCTCTCGTCCGAC 2928
DB 1329 GGGCACTGCTAGACGGAGTACAGAGCCACAAAGGTGCGAAGGAGACAAGCTGG 1388
QY 2929 CAGCTGCAAGAGATCAAGACGCTCGCCCATCCCGAGGCGCAAGAGATCTCCAAAGGA 2988
DB 1389 CAGAGACAAAAGCAGATTTCAGACCTCCCGGTGCTGTCCGCAAGATAGC----- 1440
QY 2989 GAAGGAGGAGCGCAAGGATTCGCGATCAGCAGTACTCGCGCAAGCCTTCAGATT 3048
DB 1441 -AGAGATGAAGAGAGAGAGTCTCGATCAAGAGTTACTTCCCTTATGCTTTCAATT 1499
QY 3049 CTTTCATGAGCAGACATAGAGACGTGATCAAGTCTGATCGCCAGCGCAGTATCGCAA 3108
DB 1500 CTTTCATGAGCAACCTGGAGATGTCTCAAAACCTACCAGCAGAGGTTCAGCCGGAG 1559
QY 3109 GAATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATTCGAGAT 3168
DB 1560 GCTACAGCTGGAGAGGAAATGGCAAGCTCGGCTCTGTGAGGCCGAGCAGGAGCAGAT 1619
QY 3169 GAGGAAATGCTGAACAAAGAGAGCAACTACATTCGATTGAGCGCGCCAGATGGA 3228
DB 1620 GAGGAAATGCTTACCAGAGAGAGTCTAACTACACCGGCTGAAGGGCGCAAGATGGA 1679
QY 3229 CAAGAGATGTTGCTCAAACTGAAGCCCATTTGGAGTGGGTGCTTATTTGGCGAGGTAAAGCT 3288
DB 1680 CAAGTCCATGTTGTGAATCAAGACTCTAGGCATCGGTGCCTTTGGGGAAGTGTGCTT 1739

QY 3289 GGTGAGCAAAATCGATACCTCGAACCACTTTGTATGCGATGAAACCCCTGCGGAAGACGGA 3348
DB 1740 CGCTGTTAAGCTGGA---CACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGGA 1796
QY 3349 CGTTCTCAAGCGGATCAGGTGGGACACGCTGAAGCCCGAGAGGGATATCTTCGCGGAAGC 3408
DB 1797 TGTCTCAAGCGGATCAGGTGGGACATGTTCAAGGCTCAGAGGGACATCTCTGGTGAAGC 1856
QY 3409 CGACAATAACTGGGTGAAGTTGTACTACAGCTTCCAGGACAAAGGATAATCTCTACTTT 3468
DB 1857 AGACAATGAGTGGGTGTCAAACTTACTACTCTCTCCAGGAAGACACCTGTACTTT 1916
QY 3469 TGTGATGACTATACATACAGGTGGTGTCTGATGTCTGCTCATCAAACTGGGCATTTT 3528
DB 1917 TGTGATGACTATACATACAGGTGGGATATGATGAGCCTGCTGATCAGGATGGAGTCTT 1976
QY 3529 CGAGGAACTGGCCAGATTTCTATACCTCGCGAGGTACCTGCGCGCTGGACAGGTTCA 3588
DB 1977 CCCTGAGCACCTGGCCGCTTCTATCTGAGAGTTGACCTGGCCATTTGAAGTGTCCA 2036
QY 3589 CAAAATGGCTTCTATCACAGAGACATCAAGCCTGACAACTACTCTATCTGATAGGAGCGG 3648
DB 2037 CAAGATGGCTTTATCCACCGGACATCAAGCCTGACAACTACTCTACGACTGGATGG 2096
QY 3649 ACACATAAGCTCACCGACTTTGGCCTGTGCACGGGATTCGATGGAGCGCAACTCGAA 3708
DB 2097 TCATATTAAGCTGACAGATTTTGGCCTGTGACCTGGATTCAGGTGGAATCAATTCAA 2156
QY 3709 GTACTACAGAGAAACGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAGGAAT- 3767
DB 2157 GTACTACC---AGAAAGGGAACCACTGAGACAGGACAGCATGGAGCCGCTGACCTGTG 2213
QY 3768 -----ACTCCGAGAACGCGACCAAGCCGCTGCTGGAGGCGGCACCGAT 3813
DB 2214 GGACGATGTTTCCAACTGTCTGCTGTGGAGACAGGTTAAGACCTTGGAGCAGAGGGCGCA 2273
QY 3814 GCGCATCACAAAGAGTCTTGGCCCACTCGCTGTGGGACACCCCGAATCATAGTCTCC 3873
DB 2274 GAAGCAGCACCAGAGGTGCTTGGCAGCATCTCTTGTGGGACACCAAAATTCATCGCTCC 2333
QY 3874 CGAGGTCTGGAGAGAGTGGGTACACGACGCTGTGAGCTACTGGAGCGTGGCGCTCAT 3933
DB 2334 GGAGGTCTTCTCGCAAGAGGTATACGACGCTCTGTGACTGGTGGAGCGCTGCTGTGAT 2393
QY 3934 CCTTATGAGATGCTGTGGGTGAGCGCCCTTCTGCGCAACAGTCCGCTGGGAAGACGA 3993
DB 2394 TCTCTTTGAGATGCTGTGGGAGCGCCCTTCTTGGCCCCACCCCGACAGAGACGCA 2453
QY 3994 ACAAAAGTTCATCACTGGGAGAAACCTGCAATATTCCGCGCGAGCGCGAGTTATCCCG 4053
DB 2454 GCTGAAGGTGATCACTGGGAGAGACGCTGCATATCCCTACGAGGTGAGGCTCAGCGC 2513
QY 4054 CGAGGCTACGAGTTGATAGGAGGCTCTGTGCTCGCTGCGCTGACAAAGCGCTGGCGA 4110
DB 2514 TGAGGCGCGAGACCTATCAAGAGCTGTGCTCGCGGCTGACTGCGCGCTGGCGAGGA 2573
QY 4111 GAGCGCTGAGCAGGTCAAGAGCCACAGCTTCTTCAAGGGCATCGACTTTGC---GGACAT 4167
DB 2574 TGGGCGAGATGACCTCAAGGCACACCCGCTTCTTCAACACCATCGACTTTTCCCGTGCAT 2633
QY 4168 GCGGAAGCAGAAAGCGCCCTTACATACCGGAAATCAAGCAACCCACGACACATCAACTT 4227
DB 2634 CCGAAAGCGCTGCACCTTACGCTTACGCTCCACCATCAGCCACCCCATGACACCTTCAATTT 2693
QY 4228 TGATCCCGTGGATCCGAGAGCTGGCTCGATGACTCCACCATGAGCAGCGCGCATGA 4287
DB 2694 TGACCCCGTGGATGAAGAAAGCCCTTGGCACGAGGCGCGGAGAGAGAGCCCAAGGCGCTG 2753
QY 4288 TGTGACCAAGATGACCGCACATTTTC-----CAGGGCTTTTTCGAATTTACCTT 4335
DB 2754 GGACAGCTGGCTCCCGCAGCAGCAGCATCCAGACGACGCTTCTATGATGTTACCTT 2813
QY 4336 CCGTGGCTTCTTCGACAGCAA 4356

```

Db 2814 CCGCAGGTTCTTCGATGACAA 2834
|||||
AB0233958
Mus musculus MmLATS2 mRNA for warts/lats-like kinase, complete cds.
ACCESSION AB0233958
VERSION AB0233958.1 GI:7212785
KEYWORDS MmLATS2; warts/lats-like kinase.
SOURCE Mus musculus cDNA TO mRNA.
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Yabuta,N., Fujii,T., Copeland,N.G., Gilbert,D.J., Jenkins,N.A.,
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Nishiguchi,H., Endo,Y., Toji,S., Tanaka,H., Nishimune,Y. and
Nojima,H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian
homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
REFERENCE 2 (bases 1 to 3460)
AUTHORS Nojima,H. and Fujii,T.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for
Microbial Diseases, Osaka University, Department of Molecular
Genetics; Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980,
Fax:81-6-6875-5192)
COMMENT Sequence updated (06-Jan-2000).
FEATURES
Location/Qualifiers
1..3460
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3460
/gene="MmLATS2"
134..3262
/gene="MmLATS2"
/feature="WALR"
/codon_start=1
/product="warts/lats-like kinase"
/protein_id="BAA92380.1"
/db_xref="GI:7212786"
/translation="MRPKTEPATYSGNSRQRLQRIEGLKOPSKASTOGLLVGPSND
TSLDAKVLGSDASRQQRATPKGYPQKALREIRYSLLPANESGTSAAEYVRQM
LOELVAGCDQEMAGRLQTSRSIEAALEYISKMGYLDPRNEQIVRVIKOTSPKG
LAPTVTRPFEFGTEALPSYHRLGGANYEGPALEEMPROYLDLFPAGAGTHGA
QAQHPKPGYSTAVPSAHFPPTHYGRHLLSEQPGYGVQRSSFPQNTPPDAYSSMA
KAQGGPPLTFPAHAGLYTASHHKPAATPPGAHPLHLVLTGRPTFTGESSAOAVLAP
SNSLNADLYELGSPVPSAAPLABRDSLOKQGLEASRPHVAFRAGPSRTSNFNPO
EPSPALNTVTAATHLHPVKSVRVLRPEQTAVGFSHPANVAAPATAPATESLETK
EGSAGPHLDVDYGGSERCPPPYKHLHLLPSKSEQYSDVLDLSCTVQOQSLRGTE
QDRSKHGAKGRADKKQIQTSVPVVRKNSDEERKRSIKSYSPYAFKPFMEQ
HVENYIKYQKRSRLQLEQVAKAGLCEAQEQMRKLYOKESYNRLKRAKMDKS
MFVTKTIGAGFQSVCLACKLDLHALVAMKTLRKQDVLNRNOVAHVAERDILAEAD
NEWVKLYSGKDSLVYVADYIPCGDMMSLLIRNEVEPEHLAFYIAELTLAESV
HKMGFIHRDIPDNILDLGHKLTDFGLCTGFRWTHNSKIYQKGNHQRNMEFPG
LWDDVNSRCGRDLTLEORAQKHORCLASLVTGPNYIAPEVLRLKGYTLQCDWS
VGVLIFLNVGPPPLFAPTEQTLKVNWSTLHPTQVRLSARLDLITKLCAAD
CRLGRDADMDLKAHPFNFTIDFSRDKOPAPYVPTISHPMDTSNFDVDESPWHEA
SGESAKMDTLASPSKHPHAHYEFTFRFRFDNDNGYFPCPKPEPAESADPGDADL
EGAESGCPQVYV"
3460
polyA_site
BASE COUNT 861 a 992 c 963 g 644 t
ORIGIN
Query Match 9.9%; Score 565.4; DB 10; Length 3460;
Best Local Similarity 65.3%; Pred. No. 9.8e-101;
Matches 954; Conservative 4; Mismatches 455; Indels 48; Gaps 7;

```


QY	3304	TAC	TCGAA	CCATT	TG	TAT	CGG	AT	GA	AA	CC	CT	CGG	AA	AG	CGG	AG	T	CT	CA	AG	CGG	AA	3363					
DB	2441	--	CAC	T	CGC	CC	TG	T	AC	CC	CA	T	GA	AC	CC	T	AGG	AA	A	AGG	AT	G	T	CT	CA	AG	CGG	AA	2497
QY	3364	TC	AGG	T	GC	CA	CG	T	GA	AG	CC	CG	G	AG	G	G	AT	A	T	C	T	C	G	G	A	A	A	A	3423
DB	2498	TC	AGG	T	GC	CA	CG	T	GA	AG	CC	CG	G	AG	G	AC	A	T	C	T	C	G	G	A	A	A	A	A	3497
QY	3424	GG	T	GA	AG	T	GC	T	TA	CA	AG	CT	T	T	C	AG	G	A	A	G	A	T	A	T	C	T	G	T	2557
DB	2558	GG	T	CA	AA	CT	C	T	A	C	T	T	C	CA	G	A	AA	G	A	G	A	C	T	T	T	G	T	G	3483
QY	3484	AC	CA	AG	T	G	T	G	A	T	C	T	G	A	T	C	G	C	T	A	T	C	A	A	A	T	T	T	3543
DB	2618	CC	T	G	T	G	G	G	A	C	A	T	G	A	T	C	G	A	T	G	A	T	C	G	A	T	G	A	2677
QY	3544	CA	GA	T	T	C	A	T	C	G	C	G	G	A	G	T	C	A	C	T	G	C	G	C	G	T	T	C	3603
DB	2678	CC	GG	T	T	C	A	T	C	G	C	G	G	A	G	T	C	A	C	T	G	C	G	C	G	T	T	C	2737
QY	3604	TC	AC	G	A	G	A	C	A	T	C	A	G	C	T	G	A	A	C	A	T	A	C	T	C	G	A	T	3663
DB	2738	CC	AC	C	G	A	G	A	C	A	T	C	A	G	C	T	G	A	A	C	A	T	A	C	T	C	G	A	2797
QY	3664	CG	A	T	T	T	G	C	T	G	C	A	G	G	A	T	C	C	G	A	T	C	C	G	A	T	C	A	3722
DB	2798	AG	A	T	T	G	C	C	T	C	G	A	T	C	C	G	A	T	C	C	G	A	T	C	C	G	A	T	2857
QY	3723	----	AC	G	CA	A	T	C	A	T	C	G	C	C	A	G	A	C	T	C	G	A	T	C	G	A	T	C	3774
DB	2858	GA	GC	A	T	G	T	C	A	G	A	C	A	G	A	C	A	G	A	C	A	G	A	C	A	G	A	C	2917
QY	3775	---	GA	AC	G	CA	C	CG	A	CC	C	CG	T	G	G	A	G	CG	A	C	T	G	G	A	T	G	T	3831	
DB	2918	T	C	G	G	T	G	G	G	A	C	A	G	C	T	G	A	G	A	C	C	T	G	A	G	A	C	A	2977
QY	3832	C	T	G	C	C	A	C	T	C	G	T	G	G	C	A	C	C	C	G	A	C	A	T	C	A	T	C	3891
DB	2978	C	T	G	C	C	A	T	C	A	T	C	G	T	G	G	A	C	C	C	G	A	C	A	T	C	A	3037	
QY	3892	T	G	G	T	A	C	G	A	C	T	G	C	A	T	C	T	G	A	G	G	T	G	G	T	C	A	3951	
DB	3038	A	G	G	T	A	C	A	T	C	A	T	C	G	T	G	A	G	T	G	A	G	T	G	A	T	C	3097	
QY	3952	G	G	G	T	C	A	G	C	C	C	T	T	C	G	C	A	A	C	T	C	G	C	T	G	A	A	4011	
DB	3098	G	G	G	C	A	G	C	C	C	T	T	T	T	G	C	A	C	T	C	C	C	A	G	A	A	C	3157	
QY	4012	G	G	A	A	A	C	T	G	A	T	T	C	G	C	G	A	C	C	G									

[illegible]

Brownstein,D.G., Parlow,A.F., McGrath,J. and Xu,T.
Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian
tumours and pituitary dysfunction
Nat. Genet. 21 (2), 182-186 (1999)
99140768
9988269
4 (bases 1 to 3213)
Tao, W. and Xu, T.
Direct Submission
TITLE
Submitted (04-NOV-1998) Genetics, HHMI, Yale University, 295
Congress Ave BCM 254D, New Haven, CT 06536, USA
JOURNAL
Location/Qualifiers
FEATURES
source
1. 3213
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="newborn"
<1..3213
/gene="Lats1"
<1..2889
/note="Lats1"
/gene="putative serine threonine kinase; (MUM) Lats1"
/codon_start=1
/product="large tumor suppressor 1"
/protein_id="AADI6883.1"
/db_xref="GI:4324436"
/translation="VOHSINRKQSWKSGKESLVQRHGPSLGENVVYRSESPNSQADV
GRLPSGGTAAGAAQAHNSQGVNPPQVRSVTPPPRGQTQPPRCQTQPPSPSWE
PSSQTRKYRSEMEYVLSRISPPVPGAWQEGYPPPLTSPMPNPPSQARLISVVPGR
QPLIMTOSKFNTPRGVPGVGGQSDFIHVONVTPGTSVTQPPPPYPLTFANGQSP
SALQTAAPQAFSPFANGVPSQMVNPNHSHNELYNINPGLQTAMPOSSAPQAQSS
PSGGHEIPATNPINPVSFNSNPLGSRASHANSQSATVTAITPAITQGPVKMSR
VLKPELOALATAPHSWMPQVOTVQTPFSEGTASSVPVPPVAEAPSVQGPVPPY
KLLHONPVPVSVKSPKQDEPQLPKEDSEKSSADSGDSEKKEKQITTSPIVPR
KNKKDERRESRLQSPQAFEFMEQHVENVLKSQORLHRKKOLENEMMRVQLSQD
AQDMRMLCKESNTIRUKRAMKMSMEVKIKLIGAFGEVCLRAKVDITKALYATK
TURKDDVLLRNQVHAHYKADKILAEADNMEVRLYYSFQDKNLIVPMYDIPEGDMMS
LITRMGIFPENLARFYIAELTCAASVHKMGFTHRDKPDNILIDLRGHIKLTDFGLC
TFGRWTHDSKYIQSGDHQSDMSDFSNWGDPSNCRGDLKPLERRAARHQRCHLA
TSJHNPQAKLSEASDLIIKLCPGDEBLRGKGADEIKAHPEFKITPLESSDLRQSSA
STYPIKTHFTSDNFDPVDPDKLWSDGSEENISDTLSGWYKNGKRIPEHAFYEFTFR
FFDNGNCPYKPIPEYEHISGQSEQSDDEDDQHTSSDGNRDLVYV"
BASE COUNT 946 a 764 c 696 g 807 t
ORIGIN
Query Match 9.1%; Score 522.8; DB 10; Length 3213;
Best Local Similarity 65.3%; Pred. No. 2.4e-92;
Matches 850; Conservative 4; Mismatches 42; Indels 27; Gaps 5;
QY 2984 AAGGAGAGGAGGAGGAGCGCAAGGAGTTCGGCATCAGGCAGTACTCGCCGCAAGCCCTC 3043
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1354 AACAGAAAGATGAGAACCAAGAGCTCTCGGATTCAGAGTTACTCCCCACAGGCCTTT 1413
QY 3044 AAGTCTCTATGAGCAGCAGCATAGAGAAGCTGATCAAGTCGTATCGCCAGCGCACGTAT 3103
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1414 AAGTCTCTATGAGCAGCAGCTAGAGAAGCTCCTGAAGTCTCATCAGCAGCGCTCTGCAT 1473
QY 3104 CGCAAGATCAGCTGGAGAGGAGATGCCAANAAGTGGGACTGCCGNATCAGACCCCAATC 3163
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1474 CGAAGAAGCAGCTAGAAAAATGAATGATCGGGGTTGGATTAATCTCAAGATGCCAGGAT 1533
QY 3164 GAGATGAGGAAATGCTGAACCAAGAGGAGCAACTACATTTCGATTGAAGCGCGCAAG 3223
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1534 CAATGAGAAAGATGCTTTCGCCAAGAGATCTAACTATATTCTGCTTAAAGGGCTTAA 1593
QY 3224 ATGGACAAGCAGTATGCTGCTCAACTGAAGCCCATTTGGAGTGGGTGCTATTTGGCGAGGTA 3283
|| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1594 ATGGACAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGACGGTTTGTGGAAGTC 1653
QY 3284 ACGTGTGAGCAAAATCGATACCTCGAACCAATTTGTATCGGATGAACCCCTGCGGAA 3343
|| || |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1654 TGTCTAGCAAGAAGTCGATA---CTAAAGCTTTGTATGCAACAAGACTTTCGAAG 1710

Search completed: January 16, 2003, 18:58:20
Job time : 11092.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 729.592 Seconds
(without alignments)
17655.667 Million cell updates/sec.

Title: US-09-763-334-7

Perfect score: 5720

Sequence: 1 attagcagcagcgacgca.....gcctggccaattagttac 5720

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	5708	99.8	5720	21	AAZ51508 Drosophila melanog.
2	5693.6	99.5	5720	17	AAZ42117 Drosophila melanog.
3	3628	63.4	11187	23	ABL03168 Drosophila melanog.
4	3242.2	56.7	3319	23	ABL03169 Drosophila melanog.
5	462.8	9.9	3155	21	AAZ42120 M-lats2 gene encod
6	567.8	9.9	3155	21	AAZ51507 Mouse Lats2 (large
7	558.6	9.8	5276	20	AAZ87397 Human WART1 cDNA.
8	558.6	9.8	5486	21	AAA59129 DNA encoding a tum
9	558.6	9.8	5486	21	AAA59130 DNA encoding a tum

10	557	9.7	1961	21	AAZ61160	CDNA SSI771A encod
11	557	9.7	3533	22	AAZ17165	Human cancer-inhib
12	536	9.4	1498	21	AAZ61158	CDNA SSI771A encod
13	535.2	9.4	1912	23	ABK43457	DNA encoding novel
14	522.8	9.1	3213	17	AAZ42119	M-lats gene encod
15	522.8	9.1	3213	21	AAZ51506	Mouse Lats (large
16	487.8	8.5	2043	22	AAZ03989	Human protein tyro
17	470.6	8.2	3984	17	AAZ42118	H-lats gene encod
18	470.6	8.2	3984	21	AAZ51505	Human Lats (large
19	470.6	8.2	7382	20	AAZ87396	Human WART1 cDNA.
20	469	8.2	2442	20	AAZ32981	Human warts protel
21	455.8	8.0	1374	20	AAZ32982	Human warts gene f
22	358.6	6.3	1357	22	ABA08740	Human large tumour
23	340.4	6.0	1501	22	AAZ75341	Human TGF-beta rec
24	321.4	5.6	676	22	AAZ27179	CDNA encoding nove
25	321.4	5.6	676	23	ABK43776	DNA encoding novel
26	290.8	5.1	2255	23	ABL29695	Drosophila melanog
27	286.8	5.0	2765	23	ABL02442	Drosophila melanog
28	273.6	4.8	638	21	AAZ79942	Human colon cancer
29	261.2	4.6	2101	17	AAZ31451	Drosophila Ndr ser
30	254.8	4.5	2160	22	AAZ66703	A. gossypii AG007
31	232.6	4.1	3018	17	AAZ31452	Human Ndr serine/t
32	232.6	4.1	3018	24	ABK84773	Human cDNA differe
33	232.6	4.1	3800	22	ABA08668	Human NDR homologu
34	216.6	3.8	1935	20	AAZ06834	Disease associated
35	215.4	3.8	1710	21	AAZ43403	Arabidopsis thalia
36	215.4	3.8	2003	21	AAZ43403	Arabidopsis thalia
37	215	3.8	4983	22	AAZ44629	Novel protein kina
38	204.2	3.6	1894	21	AAZ39567	Arabidopsis thalia
39	204.2	3.6	2001	21	AAZ36174	Arabidopsis thalia
40	202.6	3.5	1922	21	AAZ49426	Arabidopsis thalia
41	164.6	2.9	678	21	AAZ79982	Human colon cancer
42	161.6	2.8	734	23	ABK43709	DNA encoding novel
43	161.2	2.8	734	22	AAZ27161	CDNA encoding nove
44	161.2	2.8	734	22	AAZ56728	Human immune/haema
45	161.2	2.8	734	23	ABK43985	DNA encoding novel

ALIGNMENTS

RESULT 1

AAZ51508

ID AAZ51508 standard; DNA; 5720 BP.

XX AAZ51508;

AC AAZ51508;

XX 21-JUN-2000 (first entry)

DT Drosophila melanogaster Lats (large tumour suppressor) DNA.

DE Fruit fly; Lats; large tumour suppressor; cytostatic; vulnary;
KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
KW treatment; prevention; screening; cancer; skin; ovarian tumour;
KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
KW LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Drosophila melanogaster.

OS Drosophila melanogaster.

XX Key Location/Qualifiers

FT CDS 1103..4402

FT /*tag= a

FT /product= "Lats protein"

FT polyA_signal 4655..4660

FT /*tag= b

FT misc_feature 5013..5142

FT /*tag= c

FT /note= "this region is identical to the 1-141

FT nucleotides of Drosophila plc-21 transcript"

XX WO200010602-A1.

PN

XX PD 02-MAR-2000. 18-AUG-1999; 99WO-US19068. 18-AUG-1998; 98US-0096996. 18-AUG-1998; 98US-0096997. (UYVA) UNIV YALE.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S; Turenchalk GS, Stewart RA; WPI: 2000-246496/21. P-PSDB; AAY70393.

XX Use of lats proteins, complexes of lats and cdc2 for treating cancer that is refractory to treatment by standard chemotherapy and radiation therapy, and disorders associated with aberrant levels of cdc2 activity

XX Claim 44; Fig 15; 134pp; English.

XX The present sequence is a DNA encoding Drosophila lats (large tumour suppressor) protein which is a cell overproliferation inhibitor and a negative regulator of cell cycle-dependent kinase cdc2/cyclin A.

XX The present sequence is useful for treating cancer that is refractory to standard chemotherapy or radiation therapy such as hyperplasia, metaplasia, or dysplasia, and disorders associated with aberrant levels of cdc2 activity. Conditions treated by promoting cdc2 function include degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds. An animal model preferably a mouse, in which a lats gene has been disrupted by homologous recombination, e.g. a lats knock-out mouse, is used for screening compounds that can be used to treat or prevent cancer, particularly skin cancer, soft tissue sarcomas and ovarian tumours, and disorders associated with pituitary dysfunction e.g. luteinizing hormone (LH) hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

XX Sequence 5720 BP; 1684 A; 1491 C; 1457 G; 1088 T; 0 other;

Query Match 99.8%; Score 5708; DB 21; Length 5720; Best Local Similarity 99.5%; Pred. No.0; Matches 5693; Conservative 26; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTAGCAGCGGCGGAGCAACAAACACGAAATTAATTTTACTTAATTAAGCCAAAGCC 60

DB 1 ATCTAGCAGCGGCGGAGCAACAAACACGAAATTAATTTTACTTAATTAAGCCAAAGCC 60

QY 61 GCATCGGAATGCCCTGAAATCGGATTCGAATGCACGCGGAAAGTGGTGGCAAGC 120

DB 61 GCATCGGAATGCCCTGAAATCGGATTCGAATGCACGCGGAAAGTGGTGGCAAGC 120

QY 121 GAGTGAATCAAGTGAATACGTCGGCAATATCAGCGAATTCGTCGCAAGCAAGGA 180

DB 121 GAGTGAATCAAGTGAATACGTCGGCAATATCAGCGAATTCGTCGCAAGCAAGGA 180

QY 181 AAACGGAGAAAAGAGGAAAGCAATTAAGTGCCTGTGGGAAACGCGGAAAGCGGA 240

DB 181 AAACGGAGAAAAGAGGAAAGCAATTAAGTGCCTGTGGGAAACGCGGAAAGCGGA 240

QY 241 GAACAAAGAGCGGAAAGAGGAAATTCGTCGCAAGCAATTCGTCGCAAGCAAGGA 300

DB 241 GAACAAAGAGCGGAAAGAGGAAATTCGTCGCAAGCAATTCGTCGCAAGCAAGGA 300

QY 301 AGCTCCAAAGTTGGCGGCATCGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 360

DB 301 AGCTCCAAAGTTGGCGGCATCGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 360

QY 361 GAATCGGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420

DB 361 GAATCGGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420

QY 421 ATTCAATCATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

DB 421 ATTCAATCATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

QY 481 GTCTCTCGGCTCGGCGAGCAATCTCTGATAAGCAGGAGGAAATTAATTAATTAATTAATTA 540

DB 481 GTCTCTCGGCTCGGCGAGCAATCTCTGATAAGCAGGAGGAAATTAATTAATTAATTAATTA 540

QY 541 AATACAAATACAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 600

DB 541 AATACAAATACAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 600

QY 601 GCGAAAGGGGGGTTCTCTTATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 660

DB 601 GCGAAAGGGGGGTTCTCTTATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 660

QY 661 GTGGAAGAAATTCGCGCGGAGTGAATAATTCGCGCGGAGTGAATAATTCGCGCGGAGTGAATA 720

DB 661 GTGGAAGAAATTCGCGCGGAGTGAATAATTCGCGCGGAGTGAATAATTCGCGCGGAGTGAATA 720

QY 721 ACCGCGAGAAAGAGCGGAGCAACACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780

DB 721 ACCGCGAGAAAGAGCGGAGCAACACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780

QY 781 GAGATAATAAGCATCTCTTAAACCCGCTTAGTTCTGTTAGTCTCGCCACRGATATA 840

DB 781 GAGATAATAAGCATCTCTTAAACCCGCTTAGTTCTGTTAGTCTCGCCACRGATATA 840

QY 841 GATATTCAAAGCAAAAGTGGTCTCGGCGATCGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGT 900

DB 841 GATATTCAAAGCAAAAGTGGTCTCGGCGATCGGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGT 900

QY 901 TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 960

DB 901 TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 960

QY 961 CAAATCAATTCGCGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020

DB 961 CAAATCAATTCGCGCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020

QY 1021 TGTCTGT 1080

DB 1021 TGTCTGT 1080

QY 1081 CACAGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1140

DB 1081 CACAGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1140

QY 1141 TGATAAATACACGCGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1200

DB 1141 TGATAAATACACGCGGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1200

QY 1201 ACAAAATACCATAGGAATTAATCAGAAATACACACCTCTCGGATACACGCGGAGCAGCAGCAGCAG 1260

DB 1201 ACAAAATACCATAGGAATTAATCAGAAATACACACCTCTCGGATACACGCGGAGCAGCAGCAGCAG 1260

QY 1261 ACCCAAGATGCACTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

DB 1261 ACCCAAGATGCACTTACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320

QY 1321 CTCGCGCT 1380

DB 1321 CTCGCGCT 1380

QY 1381 CGAGCGCGCT 1440

DB 1381 CGAGCGCGCT 1440

QY 1441 ACCTGTSCCAAGATGATGAGCGGCTTAATGCAACAACTGATCCGCGGAGCAGCAGCAGCAGCAGCAG 1500

DB 1441 ACCTGTSCCAAGATGATGAGCGGCTTAATGCAACAACTGATCCGCGGAGCAGCAGCAGCAGCAGCAG 1500

QY 1501 CGAACGGGACAGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1560

Db 1501 CGAAGGACACGCGAGCAGTCACTACCTCGCTGAGTCCGGCTCTGGAGCTCGGAGC 1560
 Qy 1561 CGGTAGCTCCGATCGGACACCCGCCATTCGCACACACACCCAGCGGAGCTCGAGAC 1620
 Db 1561 CGGTAGCTCCGATCGGACACCCGCCATTCGCACACACACCCAGCGGAGCTCGAGAC 1620
 Qy 1621 GGTGGTAAATCCAGGTGGAATGTTGATTTCTCCGTCGCAAGCGGTTTCAGTGAAGT 1680
 Db 1621 GGTGGTAAATCCAGGTGGAATGTTGATTTCTCCGTCGCAAGCGGTTTCAGTGAAGT 1680
 Qy 1681 GGCTCCACCGCGCGCGCGCCGACGCAATTCACCGCCCTSCAGCGCGCCACGCCCCACC 1740
 Db 1681 GGCTCCACCGCGCGCGCGCCGACGCAATTCACCGCCCTSCAGCGCGCCACGCCCCACC 1740
 Qy 1741 GCCAGTCCCGCCACCGACCGAGCGGTAGTGAAGGGCGATCACGGCCCTGAACACCG 1800
 Db 1741 GCCAGTCCCGCCACCGACCGAGCGGTAGTGAAGGGCGATCACGGCCCTGAACACCG 1800
 Qy 1801 CCCCGCGCGATAGCGCCACCCACTCAGCGMGCAACTCACTGTATAACCCARAACGG 1860
 Db 1801 CCCCGCGCGATAGCGCCACCCACTCAGCGMGCAACTCACTGTATAACCCARAACGG 1860
 Qy 1861 GCTGAAGAACCCGACGACGAGTTGACGACGAGTGAAGTCCCTGAACCTATACCCAGG 1920
 Db 1861 GCTGAAGAACCCGACGACGAGTTGACGACGAGTGAAGTCCCTGAACCTATACCCAGG 1920
 Qy 1921 CGGAGGAGTGGAGCAGTGTGGAGCCACCGCGCCCTTACCTAATTCAGGCGGAGCGG 1980
 Db 1921 CGGAGGAGTGGAGCAGTGTGGAGCCACCGCGCCCTTACCTAATTCAGGCGGAGCGG 1980
 Qy 1981 AGGAGCAGCAGCGCGCGCCGACACCCAGTTACAGCGCTCCATGCACTCGCGCAGTC 2040
 Db 1981 AGGAGCAGCAGCGCGCGCCGACACCCAGTTACAGCGCTCCATGCACTCGCGCAGTC 2040
 Qy 2041 GCCCACAAATCCCAACAAATCGGACTACAGGAAATCCCGGAGCAGTGGGATATACCTCGC 2100
 Db 2041 GCCCACAAATCCCAACAAATCGGACTACAGGAAATCCCGGAGCAGTGGGATATACCTCGC 2100
 Qy 2101 CACCTCGCGGCTCGCGAGCCCATAACTGTCTGCTCGCGCGCGCGCGCTGGCGAA 2160
 Db 2101 CACCTCGCGGCTCGCGAGCCCATAACTGTCTGCTCGCGCGCGCGCGCTGGCGAA 2160
 Qy 2161 GCCCAACACAGTCTACCGAGCGGAGTCAAGCAGCGGATCATCATGAGAGTGTGAA 2220
 Db 2161 GCCCAACACAGTCTACCGAGCGGAGTCAAGCAGCGGATCATCATGAGAGTGTGAA 2220
 Qy 2221 GAGCAGCAGTCCAAAGCCGCTGCTGCAACAGCAGTGGCGGCCAATCGCCATCGAG 2280
 Db 2221 GAGCAGCAGTCCAAAGCCGCTGCTGCAACAGCAGTGGCGGCCAATCGCCATCGAG 2280
 Qy 2281 TGCCCTCGCGCAGCAATTCACAGTCCAGTGTGGCGCTCCACCTTTACCTTCAGAA 2340
 Db 2281 TGCCCTCGCGCAGCAATTCACAGTCCAGTGTGGCGCTCCACCTTTACCTTCAGAA 2340
 Qy 2341 GTCCGCGCAGTGTGACGACGACGACAGCAGCGCGCGCCACACAGCAGCAGCA 2400
 Db 2341 GTCCGCGCAGTGTGACGACGACGACAGCAGCGCGCGCCACACAGCAGCAGCA 2400
 Qy 2401 TCAGCACCAGCAATCCAAACACCAAGCCMACCAGCCCTTGGTGGTCTGAAYAG 2460
 Db 2401 TCAGCACCAGCAATCCAAACACCAAGCCMACCAGCCCTTGGTGGTCTGAAYAG 2460
 Qy 2461 CAAGCCCAATTCGCTGGAGCCACCGTCTATGCCAAGAGCATGAGGCCAAGGGCGCCAC 2520
 Db 2461 CAAGCCCAATTCGCTGGAGCCACCGTCTATGCCAAGAGCATGAGGCCAAGGGCGCCAC 2520
 Qy 2521 GGTGGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2580
 Db 2521 GGTGGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2580
 Qy 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640

Db 2581 ACAGCAGCAGCAACAGCAGCAACAGCAACTCAGAGCCTTGAGGGTGTCTCCAGGCACAGCG 2640
 Qy 2641 TCAGAGGAGCGGATCAACGGGAGCGGAAACGGGATCAGCAGAGCTGGCCACACGGA 2700
 Db 2641 TCAGAGGAGCGGATCAACGGGAGCGGAAACGGGATCAGCAGAGCTGGCCACACGGA 2700
 Qy 2701 TCCTGGCGCGGAGATGCTTCCGGCGCGCCCTATCAGAGCAACACCAACACACAGCA 2760
 Db 2701 TCCTGGCGCGGAGATGCTTCCGGCGCGCCCTATCAGAGCAACACCAACACACAGCA 2760
 Qy 2761 GATCAACCGCGGAGCTGCAACACACACATCAGATAGCAACAGCAACCTGGGAG 2820
 Db 2761 GATCAACCGCGGAGCTGCAACACACACATCAGATAGCAACAGCAACCTGGGAG 2820
 Qy 2821 GACACACCCATTCGGCTGCCAATACAATAACACTCTCTCCACACGGGCGGAATAG 2880
 Db 2821 GACACACCCATTCGGCTGCCAATACAATAACACTCTCTCCACACGGGCGGAATAG 2880
 Qy 2881 CTCGGGCGGACAGCAACGGATCCAGCGCACCGCCCTCTCTCGTGCAGCAGCTGCAAGAA 2940
 Db 2881 CTCGGGCGGACAGCAACGGATCCAGCGCACCGCCCTCTCTCGTGCAGCAGCTGCAAGAA 2940
 Qy 2941 GATCAAGCAGCCTCGCCCTATCCCGGAGCGCAAGAGATCTCCAGGAGAGGAGGAG 3000
 Db 2941 GATCAAGCAGCCTCGCCCTATCCCGGAGCGCAAGAGATCTCCAGGAGAGGAGGAG 3000
 Qy 3001 GCGCAAGGAGTCCGCATCAGCAGTACTCGCGCAAGCCTTCAAGTCTTCATGGAGCA 3060
 Db 3001 GCGCAAGGAGTCCGCATCAGCAGTACTCGCGCAAGCCTTCAAGTCTTCATGGAGCA 3060
 Qy 3061 GCACATAGAGAACGTGATCAAGTGTATCGCAGCGCACGTATCGCAAGATTCAGCTGGA 3120
 Db 3061 GCACATAGAGAACGTGATCAAGTGTATCGCAGCGCACGTATCGCAAGATTCAGCTGGA 3120
 Qy 3121 GAAGGAGATGCACAAAGTGGAGTGCCTCGATCAGACCCAAATCGAGATGAGGAAATGCT 3180
 Db 3121 GAAGGAGATGCACAAAGTGGAGTGCCTCGATCAGACCCAAATCGAGATGAGGAAATGCT 3180
 Qy 3181 GAACCAAGAGAGCAACTACATTCGATGAAGCGCGCAAGATGGACAAGAGCATGTT 3240
 Db 3181 GAACCAAGAGAGCAACTACATTCGATGAAGCGCGCAAGATGGACAAGAGCATGTT 3240
 Qy 3241 CGTCAAACTGAAGCCCTTGGAGTGGTGCATTTGGCGAGGTAACTGCTGGAGCAAAAT 3300
 Db 3241 CGTCAAACTGAAGCCCTTGGAGTGGTGCATTTGGCGAGGTAACTGCTGGAGCAAAAT 3300
 Qy 3301 CGATACCTCGAACCATTTGATGGATGAAAACCTCGCGGAGAGCGGAGTCTCAAGCG 3360
 Db 3301 CGATACCTCGAACCATTTGATGGATGAAAACCTCGCGGAGAGCGGAGTCTCAAGCG 3360
 Qy 3361 GAATCAGGTGCGACACGCTGAAGCGGAGGAGATATCTCGCGAAGCGGAGCAATACTG 3420
 Db 3361 GAATCAGGTGCGACACGCTGAAGCGGAGGAGATATCTCGCGAAGCGGAGCAATACTG 3420
 Qy 3421 GGTGGTGAAGTGTACTACAGTCTCCAGGACAGGATATCTGCTACTTTGTGATGAGTAA 3480
 Db 3421 GGTGGTGAAGTGTACTACAGTCTCCAGGACAGGATATCTGCTACTTTGTGATGAGTAA 3480
 Qy 3481 CATACAGGTGGTGTGATGCTGCTCTCATCAAACTGGGCAATTTTCGAGGAGGAACT 3540
 Db 3481 CATACAGGTGGTGTGATGCTGCTCTCATCAAACTGGGCAATTTTCGAGGAGGAACT 3540
 Qy 3541 GGCCAGATTTACATCCCGGAGGTCACTCGCGGTGACAGCTTCAAAAATGGGCTT 3600
 Db 3541 GGCCAGATTTACATCCCGGAGGTCACTCGCGGTGACAGCTTCAAAAATGGGCTT 3600
 Qy 3601 CATTCACAGACATCAAGCCTGACACATACTCATCATAGGAGCGGACACATAAAGCT 3660
 Db 3601 CATTCACAGACATCAAGCCTGACACATACTCATCATAGGAGCGGACACATAAAGCT 3660
 Qy 3661 CACCGACTTTGGCTGTGACGCGGATTCGATGAGCAGCAACTCGAAGTACTACACAGA 3720
 Db 3661 CACCGACTTTGGCTGTGACGCGGATTCGATGAGCAGCAACTCGAAGTACTACACAGA 3720

QY	4801	ACTGGAATCGAACCTTCTCGTGTCTCAAGCAAGCAAAAGCAAAACAAACGCC	4861
Db	4801	ACTGGAATCGAACCTTCTCGTGTCTCAAGCAAGCAAAAGCAAAACAAACGCC	4860
QY	4861	TTAAACTAAAYGAGACGCGAAATTACCCAAACACATTCACTCCTCTCTTCTCCACCTCC	4920
Db	4861	TTAAACTAAATGAGACGCGAAATTACCCAAACACATTCACTCCTCTCTTCTCCACCTCC	4920
QY	4921	GATCGGTGCCGGATTTCGAACCTCAGCAGGCTGGTTGTCATCGCGGCATCCCATTKACTTCC	4980
Db	4921	GATCGGTGCCGGATTTCGAACCTCAGCAGGCTGGTTGTCATCGCGGCATCCCATTKACTTCC	4980
QY	4981	CATTTCAGAAATTGAGATTCGCGAGGTGCGATGGAGAAACGACGAGACCAAAAGTCCGAC	5040
Db	4981	CATTTCAGAAATTGAGATTCGCGAGGTGCGATGGAGAAACGACGAGACCAAAAGTCCGAC	5040
QY	5041	GGCAGCGCATATAAGCGGGTCTTATAAGCCTTAATCTAAATCTTAAACTCGGGAGAACAGGACC	5100
Db	5041	GGCAGCGCATATAAGCGGGTCTTATAAGCCTTAATCTAAATCTTAAACTCGGGAGAACAGGACC	5100
QY	5101	YATGTATGTCCCTGCTATCAAAATTCGTCTCTATCAGTCTCTTCANCTGNGTACGACCCCCAC	5160
Db	5101	TATGTATGTCCCTGCTATCAAAATTCGTCTCTATCAGTCTCTTCATCTGTGTACGACCCCCAC	5160
QY	5161	CCCCCCCCCTCCCATCCAAAAGAAACAAACTTAGACGTAAGCCTATGTGAAAGCTAGCAAT	5220
Db	5161	CCCCCCCCCTCCCATCCAAAAGAAACAAACTTAGACGTAAGCCTATGTGAAAGCTAGCAAT	5220
QY	5221	GTTTAGACCAACTTGTGTGAATGCCAAATGAAATGTTTAGCCCCACGAGGAAACGCGGGG	5280
Db	5221	GTTTAGACCAACTTGTGTGAATGCCAAATGAAATGTTTAGCCCCACGAGGAAACGCGGGG	5280
QY	5281	GAAATTCACACTTATCTCTCATAGCAACCGAAAGAAAGAAAGAAAAAACAACAG	5340
Db	5281	GAAATTCACACTTATCTCTCATAGCAACCGAAAGAAAGAAAGAAAAAACAACAG	5340
QY	5341	AAACAGTACGAAAAATGTAAATCTCTTAAATGTAATTTGTAAGAACACGTTAATTGT	5400
Db	5341	AAACAGTACGAAAAATGTAAATCTCTTAAATGTAATTTGTAAGAACACGTTAATTGT	5400
QY	5401	AATCTATGCTAGAGTTGTGTAGCGCCCTAAGATGTTTTTATGTTTATAGACCGCTAACCG	5460
Db	5401	AATCTATGCTAGAGTTGTGTAGCGCCCTAAGATGTTTTTATGTTTATAGACCGCTAACCG	5460
QY	5461	TAACTAGTGTAAATTCCTAAACACTAAGCGAGAGTACAGTACATGGTTTTTTTTTGTGTC	5520
Db	5461	TAACTAGTGTAAATTCCTAAACACTAAGCGAGAGTACAGTACATGGTTTTTTTTTGTGTC	5520
QY	5521	GTAGGTTCTGTGGAAAATGCTTTAAGCGGAAACGATTTGTTTTCTCTTTAATTAGCTTCA	5580
Db	5521	GTAGGTTCTGTGGAAAATGCTTTAAGCGGAAACGATTTGTTTTCTCTTTAATTAGCTTCA	5580
QY	5581	GTTTGTATGTGCGTGTGTTTTTATATAGCTATATATAGTCCATCTGAAATATTCGTGA	5640
Db	5581	GTTTGTATGTGCGTGTGTTTTTATATAGCTATATATAGTCCATCTGAAATATTCGTGA	5640
QY	5641	TGGAGCCTATTTTAAATCTGAGATCGAGCTAATTAAGAGGAATACAAACAAACTCTGTGT	5700
Db	5641	TGGAGCCTATTTTAAATCTGAGATCGAGCTAATTAAGAGGAATACAAACAAACTCTGTGT	5700
QY	5701	GCCTTGGCCAAATTAGTTTAC	5720
Db	5701	GCCTTGGCCAAATTAGTTTAC	5720

RESULT 2				
AAT42117				
ID	AAT42117	standard;	CDNA;	5720 BP.
XX	XX			
XX	AAT42117;			
XX	XX			
DT	22-JAN-1997	(first entry)		
XX				

```

DE  XX  Lats gene encoding large tumour suppressor.
KW  XX  Lats gene; large tumour suppressor; Drosophila melanogaster;
KW  KW  fruitfly; polyadenylation site; protein-serine/threonine-kinase;
KW  KW  cell proliferation; antisense; dominant-negative; cancer;
KW  KW  degenerative disorder; trauma; growth deficiency; therapy;
KW  KW  antitumour; vulnerable; diagnostic; transgenic plant;
KW  KW  transgenic animal; growth; senescence; ds.
XX  OS  Drosophila melanogaster.
XX  XX
XX  XX  Location/Qualifiers
FH  Key 1..191
FH  misc_feature /tag= a
FT  FT /note= "Sequence from clone cDNA-9"
FT  FT 1..1102
FT  FT /tag= b
FT  FT 81
FT  FT /tag= c
FT  FT /note= "A in genomic sequence"
FT  FT 192..5720
FT  FT /tag= d
FT  FT /note= "Sequence from clone cDNA-A2"
FT  FT 302..303
FT  FT /tag= e
FT  FT /note= "Intron-1 splice site"
FT  FT 330
FT  FT /tag= f
FT  FT /note= "T in genomic sequence"
FT  FT 606..608
FT  FT /tag= g
FT  FT /note= "GGA in genomic sequence"
FT  FT 821
FT  FT /tag= h
FT  FT /note= "G in genomic sequence"
FT  FT 834
FT  FT /tag= i
FT  FT /note= "A in genomic sequence"
FT  FT 1103..4402
FT  FT /tag= j
FT  FT /product= Large tumour suppressor (lats protein)
FT  FT 1225..1226
FT  FT /tag= k
FT  FT /note= "Intron-2 splice site"
FT  FT 1371..1818
FT  FT /tag= l
FT  FT /note= "lats-al deletion"
FT  FT 1447
FT  FT /tag= m
FT  FT /note= "C in genomic sequence"
FT  FT 1719
FT  FT /tag= n
FT  FT /note= "G in genomic sequence"
FT  FT 1939
FT  FT /tag= o
FT  FT /note= "T in genomic sequence"
FT  FT 2137
FT  FT /tag= p
FT  FT /note= "C in genomic sequence"
FT  FT 2431
FT  FT /tag= q
FT  FT /note= "C in genomic sequence"
FT  FT 2458
FT  FT /tag= r
FT  FT /note= "T in genomic sequence"
FT  FT 2551
FT  FT /tag= s
FT  FT /note= "G in genomic sequence"
FT  FT 2554
FT  FT /tag= t
FT  FT /note= "G in genomic sequence"
FT  FT 2555
FT  FT /tag= u
FT  FT

FT  conflict /note= "AAC in genomic sequence"
FT  FT 2556
FT  FT /tag= v
FT  FT /note= "A in genomic sequence"
FT  FT 2593
FT  FT /tag= w
FT  FT /note= "G in genomic sequence"
FT  FT 2646
FT  FT /tag= x
FT  FT /note= "GGAGCGGATCAAC insertion in genomic DNA"
FT  FT 3488..3489
FT  FT /tag= y
FT  FT /note= "Intron-3 splice site"
FT  FT 3609..3610
FT  FT /tag= z
FT  FT /note= "Intron-4 splice site"
FT  FT 3725..3726
FT  FT /tag= aa
FT  FT /note= "Intron-5 splice site"
FT  FT 3804..3805
FT  FT /tag= ab
FT  FT /note= "Intron-6 splice site"
FT  FT 3889
FT  FT /tag= ac
FT  FT /note= "A in genomic sequence"
FT  FT 3937
FT  FT /tag= ad
FT  FT /note= "C in genomic sequence"
FT  FT 3940
FT  FT /tag= ae
FT  FT /note= "T in genomic sequence"
FT  FT 4000..4001
FT  FT /tag= af
FT  FT /note= "Intron-7 splice site"
FT  FT 4021
FT  FT /tag= ag
FT  FT /note= "T in genomic sequence"
FT  FT 4021
FT  FT /tag= ah
FT  FT /note= "C in genomic sequence"
FT  FT 4210
FT  FT /tag= ai
FT  FT /note= "G in genomic sequence"
FT  FT 4300
FT  FT /tag= aj
FT  FT /note= "C in genomic sequence"
FT  FT 4309
FT  FT /tag= ak
FT  FT /note= "C in genomic sequence"
FT  FT 4402..5720
FT  FT /tag= al
FT  FT /note= "polyA signal"
FT  FT 4654..4660
FT  FT /tag= am
FT  FT /note= "conflict"
FT  FT 4797
FT  FT /tag= an
FT  FT /note= "G in genomic sequence"
FT  FT 4828..4837
FT  FT /tag= ao
FT  FT /note= "Deleted in genomic DNA"
FT  FT 4871
FT  FT /tag= ap
FT  FT /note= "C in genomic sequence"
FT  FT 4974
FT  FT /tag= aq
FT  FT /note= "T in genomic sequence"
FT  FT 5013..5153
FT  FT /tag= ar
FT  FT /note= "Region identical to plc-21 transcript"
FT  FT 5101
FT  FT /tag= as
FT  FT /note= "C in genomic sequence"
FT  FT 5104..5114
FT  FT /tag= at
FT  FT

```


Db 1381 CGGACCGGGTCCATATCCGGTATCCGGTGTGGGGTGGAGTGTGGTGTGGCGAACGG 1440
QY 1441 ACCTGTSCCAAGATGATGAGCGCCCTAATCCCAAAACAACTGATCCGGAGCGAGCAT 1500
Db 1441 ACCTGTSCCAAGATGATGAGCGCCCTAATCCCAAAACAACTGATCCGGAGCGAGCAT 1500
QY 1501 CGAACGGGACACGGGAGCAGTCACTACCTGCGCTGCGAGTCCGGCTCTGGACTCCGGAGC 1560
Db 1501 CGAACGGGACACGGGAGCAGTCACTACCTGCGCTGCGAGTCCGGCTCTGGACTCCGGAGC 1560
QY 1561 CGGTAGTCCCGATCGGACAGCCCCCATTCGCACCAACACCCAGCCGAGCTCGAGGAC 1620
Db 1561 CGGTAGTCCCGATCGGACAGCCCCCATTCGCACCAACACCCAGCCGAGCTCGAGGAC 1620
QY 1621 GGTGGGTAAATCCAGGTGGAATGTGGATTTCTCCGTGCGCAGCGGTTTCACTGAGGT 1680
Db 1621 GGTGGGTAAATCCAGGTGGAATGTGGATTTCTCCGTGCGCAGCGGTTTCACTGAGGT 1680
QY 1681 GGTTCACCGCGCGCGCCGACCAATCCACCGCTCAGCGCGGCCACCGCCGCCACC 1740
Db 1681 GGTTCACCGCGCGCGCCGACCAATCCACCGCTCAGCGCGGCCACCGCCGCCACC 1740
QY 1741 GCCAGTCCGCCGCCACCGAGCGGTACGTGAAGCGCGGATCACCGCCCTGAAACACCG 1800
Db 1741 GCCAGTCCGCCGCCACCGAGCGGTACGTGAAGCGCGGATCACCGCCCTGAAACACCG 1800
QY 1801 CCCGCCGGGATAGCGCCACCCACTCAGCGMGCACTCACTGTATATACCCARAACGG 1860
Db 1801 CCCGCCGGGATAGCGCCACCCACTCAGCGMGCACTCACTGTATATACCCARAACGG 1860
QY 1861 GCTGAAGAACCGGAGCAGTGTGACGCGAGCAGTGAAGTCCCTGAACCTATACCCAGG 1920
Db 1861 GCTGAAGAACCGGAGCAGTGTGACGCGAGCAGTGAAGTCCCTGAACCTATACCCAGG 1920
QY 1921 CGGAGGAGTGAGCAGTGTGAGGACACCGCGCCCTACTTAATCAAGCGGAGCGCG 1980
Db 1921 CGGAGGAGTGAGCAGTGTGAGGACACCGCGCCCTACTTAATCAAGCGGAGCGCG 1980
QY 1981 AGGAGCAGCAGCGCGCGCCACCCAGTTCACAGCGCTCCATGTCAGTCCGGCGAGTC 2040
Db 1981 AGGAGCAGCAGCGCGCGCGCCACCCAGTTCACAGCGCTCCATGTCAGTCCGGCGAGTC 2040
QY 2041 GCCACACAATCCCAACATCGGACTACAGGAATCCCGAGCAGTGGGATATCTCGGC 2100
Db 2041 GCCACACAATCCCAACATCGGACTACAGGAATCCCGAGCAGTGGGATATCTCGGC 2100
QY 2101 CACTCGGGGGTTCGCGAGCGCCCATTAATGTGTCTGCTGCGCGCGCGCGCTGGCGAA 2160
Db 2101 CACTCGGGGGTTCGCGAGCGCCCATTAATGTGTCTGCTGCGCGCGCGCGCTGGCGAA 2160
QY 2161 GCCACACACAGTCTTACAGCGCCAGGACTCAGCGCCGATCATCATGAGTGTGAA 2220
Db 2161 GCCACACACAGTCTTACAGCGCCAGGACTCAGCGCCGATCATCATGAGTGTGAA 2220
QY 2221 GAGCAGCAGTCCAAAGCCGTGCTGCAAAACAGCAGTGGCGGCCCAATCGCCATCGAG 2280
Db 2221 GAGCAGCAGTCCAAAGCCGTGCTGCAAAACAGCAGTGGCGGCCCAATCGCCATCGAG 2280
QY 2281 TGCTCGGGCAGCAATTCACAGTCCAGTGTGGCGCGCTCCAGCTTTACCCCTAGAA 2340
Db 2281 TGCTCGGGCAGCAATTCACAGTCCAGTGTGGCGCGCTCCAGCTTTACCCCTAGAA 2340
QY 2341 GTCCGGGCGAGTGTGCGAGCAGCAACAGGAGCAGCGCGGCCACCCAGCAGCAGCA 2400
Db 2341 GTCCGGGCGAGTGTGCGAGCAGCAACAGGAGCAGCGCGGCCACCCAGCAGCAGCA 2400
QY 2401 TCAGCAGCAGCAATTCACAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
Db 2401 TCAGCAGCAGCAATTCACAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2460
QY 2461 CAAGCCCAATTTGCTGAGCGCCAGCTTATGCCAAGAGCATGCAGGCCAAGCGGCCAC 2520
Db 2461 CAAGCCCAATTTGCTGAGCGCCAGCTTATGCCAAGAGCATGCAGGCCAAGCGGCCAC 2520

Db 2461 CAAGCCCAATTTGCTGAGCGCCAGCTTATGCCAAGAGCATGCAGGCCAAGCGGCCAC 2520
QY 2521 GGTGGTARACAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
Db 2521 GGTGGTARACAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
QY 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
Db 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
QY 2641 TCAGAGGAGCGGGATCAACGGGAGCGGAAACGGGATCAGCAGAACTTCCAGGCTTCCAGCAGCAG 2700
Db 2641 TCAGAGGAGCGGGATCAACGGGAGCGGAAACGGGATCAGCAGAACTTCCAGGCTTCCAGCAGCAG 2700
QY 2701 TCCTGGCGCGCAGATGCTTCCGGCGCGCCCTATTCAGAGCAACAAACAAACACACCGA 2760
Db 2701 TCCTGGCGCGCAGATGCTTCCGGCGCGCCCTATTCAGAGCAACAAACAAACACACCGA 2760
QY 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAAGCAACAAACCTGGCGAC 2820
Db 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAAGCAACAAACCTGGCGAC 2820
QY 2821 GACACACCCATTCGGCTGCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2880
Db 2821 GACACACCCATTCGGCTGCCAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2880
QY 2881 CTGGGGCGGAGCAGCAGGATCCACCGGCGCCACCGCTCTCTGTCGACAGCTGCAAGAA 2940
Db 2881 CTGGGGCGGAGCAGCAGGATCCACCGGCGCCACCGCTCTCTGTCGACAGCTGCAAGAA 2940
QY 2941 GATCAAGCAGCGCTCGGCCATCCGGAGCGCAAGAGATCTCCAAGGAGAGGAGGAGGA 3000
Db 2941 GATCAAGCAGCGCTCGGCCATCCGGAGCGCAAGAGATCTCCAAGGAGAGGAGGAGGA 3000
QY 3001 GCGAAGGAGTTCGGCATCAGGAGTACTCGCGCAAGCCTTCAAGTTCTTCATGAGCA 3060
Db 3001 GCGAAGGAGTTCGGCATCAGGAGTACTCGCGCAAGCCTTCAAGTTCTTCATGAGCA 3060
QY 3061 GCACATAGACAAGTGAATCAAGTCTGATTCGCGCAGCGCAGCTATCGCAAGAACTGAGTGA 3120
Db 3061 GCACATAGACAAGTGAATCAAGTCTGATTCGCGCAGCGCAGCTATCGCAAGAACTGAGTGA 3120
QY 3121 GAAGGAGATCACAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAATGCT 3180
Db 3121 GAAGGAGATCACAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAATGCT 3180
QY 3181 GACCAAAAGAGAGCACTACATTCGATTGAAGCGCGCCCAAGATGAGCAAGAGCATGTT 3240
Db 3181 GACCAAAAGAGAGCACTACATTCGATTGAAGCGCGCCCAAGATGAGCAAGAGCATGTT 3240
QY 3241 CGTCAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAACTGCTGATGAGCAAAAT 3300
Db 3241 CGTCAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAACTGCTGATGAGCAAAAT 3300
QY 3301 CGATACCTCGAAGCATTGTTGATGCGATGAAACCCCTCGGAAAGCGGACCTTCTCAAGCG 3360
Db 3301 CGATACCTCGAAGCATTGTTGATGCGATGAAACCCCTCGGAAAGCGGACCTTCTCAAGCG 3360
QY 3361 GAATCAGGTGGCAGCATGTAAGCGCGGAGGATATCTCTCGGAGCGCCGACAACTACTG 3420
Db 3361 GAATCAGGTGGCAGCATGTAAGCGCGGAGGATATCTCTCGGAGCGCCGACAACTACTG 3420
QY 3421 GGTGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3480
Db 3421 GGTGGTGAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3480
QY 3481 CATACAGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3540
Db 3481 CATACAGGTGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3540
QY 3541 GGGCAGATTCTACATCGCGGAGGTCACTCGCGCGTGGAGCAGGCTTCAAAAATGGGCTT 3600
Db 3541 GGGCAGATTCTACATCGCGGAGGTCACTCGCGCGTGGAGCAGGCTTCAAAAATGGGCTT 3600

QY 3601 CATTACAGAGACATCAAGCCTGACAACTACTCATGATAGGGACGACACATAAAGCT 3660
DB 3601 CATTACAGAGACATCAAGCCTGACAACTACTCATGATAGGGACGACACATAAAGCT 3660
QY 3661 CACCGACTTTGGCCTGTGCACGGGATTCGGATGGACGACAACTCGAAGTACTACCGAGGA 3720
DB 3661 CACCGACTTTGGCCTGTGCACGGGATTCGGATGGACGACAACTCGAAGTACTACCGAGGA 3720
QY 3721 GAACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
DB 3721 GAACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
QY 3781 ACCGAGCCACCGCTGTGGAGAGCGGACGATCGCGATCACCAAGAGTCTCTGGCCCA 3840
DB 3781 ACCGAGCCACCGCTGTGGAGAGCGGACGATCGCGATCACCAAGAGTCTCTGGCCCA 3840
QY 3841 CTCGCTGTGGGACACCGGACCTACATAGCTCCGAGGCTGCTGGAGAGTGGGTACAC 3900
DB 3841 CTCGCTGTGGGACACCGGACCTACATAGCTCCGAGGCTGCTGGAGAGTGGGTACAC 3900
QY 3901 GCAGCTGTGGACTACTGGAGCGTGGCGCTCATCTTATGAGATGCTGGTGGGTACGCC 3960
DB 3901 GCAGCTGTGGACTACTGGAGCGTGGCGCTCATCTTATGAGATGCTGGTGGGTACGCC 3960
QY 3961 GCCCTTTCTGGCCCAACAGTCCGCTGAAACGACAAAGGTCATCAACTGGGAGAAAC 4020
DB 3961 GCCCTTTCTGGCCCAACAGTCCGCTGAAACGACAAAGGTCATCAACTGGGAGAAAC 4020
QY 4021 SCTGCAATTTCCGCGCAGCGCGAGTTATCCGCGAGGCTACGGACTTGTAAAGAGCT 4080
DB 4021 SCTGCAATTTCCGCGCAGCGCGAGTTATCCGCGAGGCTACGGACTTGTAAAGAGCT 4080
QY 4081 CTGTGCGTGGCTGACAGCGCTGGGCAAGAGCTGGACGAGTCAAGAGCCACGACTT 4140
DB 4081 CTGTGCGTGGCTGACAGCGCTGGGCAAGAGCTGGACGAGTCAAGAGCCACGACTT 4140
QY 4141 CTTCAAGGCAATCGACTTTGCGGACATCGGAAGCAGAAAGCGCCCTACATACCGGAAAT 4200
DB 4141 CTTCAAGGCAATCGACTTTGCGGACATCGGAAGCAGAAAGCGCCCTACATACCGGAAAT 4200
QY 4201 CAAGCACCACCGGACACATCCACTTTGATCCGCTGGATCCGAGAGCTGCGTCGAA 4260
DB 4201 CAAGCACCACCGGACACATCCACTTTGATCCGCTGGATCCGAGAGCTGCGTCGAA 4260
QY 4261 TGACTCCACCATGAGCAGCGCGGATGATGACGAGATGACCGGCTTCCACGGCTT 4320
DB 4261 TGACTCCACCATGAGCAGCGCGGATGATGACGAGATGACCGGCTTCCACGGCTT 4320
QY 4321 TTTGCAATTTACCTTCCGCTGCTTTCGACGACAGCAGCGCGGATATGACGGAGCA 4380
DB 4321 TTTGCAATTTACCTTCCGCTGCTTTCGACGACAGCAGCGCGGATATGACGGAGCA 4380
QY 4381 TCAGGCGCGGTTTACGCTCTGAATGAGTCTCTCATGTGCGCCCAACACACCGCCGC 4440
DB 4381 TCAGGCGCGGTTTACGCTCTGAATGAGTCTCTCATGTGCGCCCAACACACCGCCGC 4440
QY 4441 CCCCAGATCATTTGTAGTCAAACTACAAAAGGGGATGAAAGCATTGAGTGGGCTT 4500
DB 4441 CCCCAGATCATTTGTAGTCAAACTACAAAAGGGGATGAAAGCATTGAGTGGGCTT 4500
QY 4501 GCATTGTAAGGAAGCGTGGCTATAGATGAACACTATCATATATATAAATATATA 4560
DB 4501 GCATTGTAAGGAAGCGTGGCTATAGATGAACACTATCATATATATAAATATATA 4560
QY 4561 GGAGACAGTAGGGGAGCTAGCTATACATAAATAATACATATATTTGATAT 4620
DB 4561 GGAGACAGTAGGGGAGCTAGCTATACATAAATAATACATATATTTGATAT 4620
QY 4621 ATATATATATATATCCCTAGGCGATCACTGAAATAATAAAGAGCGAGGAGTAG 4680
DB 4621 ATATATATATATATCCCTAGGCGATCACTGAAATAATAAAGAGCGAGGAGTAG 4680

QY 4681 AGATGAACAGAGAGGAGCGAGTCCAGACCTTTCAGCTTTAACTGACATAGTATATCCTT 4740
DB 4681 AGATGAACAGAGAGGAGCGAGTCCAGACCTTTCAGCTTTAACTGACATAGTATATCCTT 4740
QY 4741 GTGCACCTACTACTCCACACAAATATATATTTTAAATTTGTTAGAAATCAAAAGGRCAC 4800
DB 4741 GTGCACCTACTACTCCACACAAATATATATTTTAAATTTGTTAGAAATCAAAAGGRCAC 4800
QY 4801 ACTGAAATCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
DB 4801 ACTGAAATCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
QY 4861 TTAACCTAAATGAGAGCGGAATTTACCCCAACCTTCACTCTCTCTCTCTCTCTCTCTCTCT 4920
DB 4861 TTAACCTAAATGAGAGCGGAATTTACCCCAACCTTCACTCTCTCTCTCTCTCTCTCTCTCT 4920
QY 4921 GATCGGTGGCGGATTCGAACTCAGCAGCTGTTTGCATCCGCCATCCCATTTKACTTCC 4980
DB 4921 GATCGGTGGCGGATTCGAACTCAGCAGCTGTTTGCATCCGCCATCCCATTTKACTTCC 4980
QY 4981 CATTCAAGATTTGAGATTTGCGAGTGTGCGATGGAGAACGAGAGAACGAGAACGAGAACGAG 5040
DB 4981 CATTCAAGATTTGAGATTTGCGAGTGTGCGATGGAGAACGAGAGAACGAGAACGAGAACGAG 5040
QY 5041 GGCAGCGATATAAGCGGCTTTAATAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAAAT 5100
DB 5041 GGCAGCGATATAAGCGGCTTTAATAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAAAT 5100
QY 5101 YATGATGCTGCTATCCATTCGATTCGCTATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
DB 5101 YATGATGCTGCTATCCATTCGATTCGCTATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
QY 5161 CCCCCCTTCCCATCCCAAGAACAACTTAGAGTACGCTATGTGAAAAGCTAGCAAT 5220
DB 5161 CCCCCCTTCCCATCCCAAGAACAACTTAGAGTACGCTATGTGAAAAGCTAGCAAT 5220
QY 5221 GTTAGCAACACTTTGTGAATGCCAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 5280
DB 5221 GTTAGCAACACTTTGTGAATGCCAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 5280
QY 5281 GAAATTCACACTTTTCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
DB 5281 GAAATTCACACTTTTCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
QY 5341 AAACAGTACGAGAAATTTGTAATCTTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 5400
DB 5341 AAACAGTACGAGAAATTTGTAATCTTAAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 5400
QY 5401 AATCTATGCTAGAGTTGTGTAGCGCCCTAAGATGTTTTTTTATAGTATAGACCGCTAACCG 5460
DB 5401 AATCTATGCTAGAGTTGTGTAGCGCCCTAAGATGTTTTTTTATAGTATAGACCGCTAACCG 5460
QY 5461 TAATCTAGTTTAAATCCTAACACTAGCGAGTACAGTACATTTGTTTGTGTTTGTGTTGTTGTC 5520
DB 5461 TAATCTAGTTTAAATCCTAACACTAGCGAGTACAGTACATTTGTTTGTGTTTGTGTTGTTGTC 5520
QY 5521 GTAGTTCGTTGGAATACTTAACGGAACGATTTGTTTCTCTTTTAAATAGCTTCA 5580
DB 5521 GTAGTTCGTTGGAATACTTAACGGAACGATTTGTTTCTCTTTTAAATAGCTTCA 5580
QY 5581 GTTGTATGTGCGTGTGTTTTTATATGACTATATATAGTCCATCTGAATTTCTGTTG 5640
DB 5581 GTTGTATGTGCGTGTGTTTTTATATGACTATATATAGTCCATCTGAATTTCTGTTG 5640
QY 5641 TGGAGCCTTATTTAAATGTGAGTTCGAGCTTAATGAAGGAAATACAAACAACTCTGTGT 5700
DB 5641 TGGAGCCTTATTTAAATGTGAGTTCGAGCTTAATGAAGGAAATACAAACAACTCTGTGT 5700
QY 5701 GCCTTGGCCAAATAGTTTAC 5720
DB 5701 GCCTTGGCCAAATAGTTTAC 5720

Query Match	Best Local Similarity	Mismatches	Conservative	Indels	Gaps
1223	CAGAATTACACACCTCTGGGATACACGGGACCAACGAGCAGCAACGATGCACTTACTCTCT 1282				
4519	CAGAATTACACACCTCTGGGATACACGGGACCAACGAGCAGCAACGATGCACTTACTCTCT 4460				
1283	GACTATACACGCGCAAGCAGCGATGAGCGCGCCACCCCTCCCTGCTCCGCGAC 1342				
4459	GACTATACACGCGCAAGCAGCGATGAGCGCGCCACCCCTCCCTGCTCCGCGAC 4400				
1343	GTGGTATACACGCGCGCGCCCTTGTAGTCAGCCGCGCGGCTCCATATCCGTA 1402				
4399	GTGGTATACACGCGCGCGCGCCCTTGTAGTCAGCCGCGCGGCTCCATATCCGTA 4340				
1403	TCGGGTGTGGCGCTTGGAGTGTGGGTGTGGGCAACGAGCTGTSCAAAGATGATGAGC 1462				
4339	TCGGGTGTGGCGCTTGGAGTGTGGGTGTGGGCAACGAGCTGTGCAAGATGATGAGC 4280				
1463	GCCTTAATGCCAAACAACTGATCCGGAAAGCCGAGATCCGAACGGGACACGCGGACGT 1522				

QY	2600	CAACAGCAACTGCAGGCCTTTGAGGGTGCTCCAGGCACAGGCTCAGA-----	2645
Db	3139	CAACAGCAACTGCAGGCCTTTGAGGGTGCTCCAGGCACAGGCTCAGAGGAGCGGGATCAA	3080
QY	2646	---GGGAGCGGGATCAACGGGAGCGGGAAACGGGATCAGCAGAACTGGCCAAAGGAATCCT	2704
Db	3079	CGGGAGCGGGATCAACGGGAGCGGGAAACGGGATCAGCAGAACTGGCCAAAGGAATCCT	3020
QY	2705	GGCGCGCAGATGCTTTCGCCGCCGCCCTATCAGAGCAACAACAACAACAGCAGGATC	2764
Db	3019	GGCGCGCAGATGCTTTCGCCGCCGCCCTATCAGAGCAACAACAACAACAGCAGGATC	2960
QY	2765	AAACGCCGGAGCTGCACAACAACAACAATACAGATAAGCAACAGCAACCTGGCAGCACA	2824
Db	2959	AAACGCCGGAGCTGCACAACAACAACAATACAGATAAGCAACAGCAACCTGGCAGCACA	2900
QY	2825	CAACCCATTCGCCCTGCCAAATACAATAACAACCTCTCCAACACGGGCGCAATAGTCG	2884
Db	2859	CCACCCATTCGCCCTGCCAAATACAATAACAACCTCTCCAACACGGGCGCAATAGTCG	2840
QY	2885	GGCGCGCAGCAACGGATCCACCGGCACCCACCGCTCTCTGTCGACAGCTGCAAGAAGATC	2944
Db	2839	GGCGCGCAGCAACGGATCCACCGGCACCCACCGCTCTCTGTCGACAGCTGCAAGAAGATC	2780
QY	2945	AAGCACCCCTCGCCCATCCCGGAGCGCAAGAGATCTCCAAGGAGAGGAGGAGCGC	3004
Db	2779	AAGCACCCCTCGCCCATCCCGGAGCGCAAGAGATCTCCAAGGAGAGGAGGAGGAGCGC	2720
QY	3005	AAGGAGTTCGCGATCAGGCAAGTACTCGCGCAAGGCTTCAAGTTCTTCATGGAGCAGCAC	3064
Db	2719	AAGGAGTTCGCGATCAGGCAAGTACTCGCGCAAGGCTTCAAGTTCTTCATGGAGCAGCAC	2660
QY	3065	ATAGAGAACCTGATCAAGTCGTATGCCACGCCACGATATCGCAAGAATACAGCTGGAGAAG	3124
Db	2659	ATAGAGAACCTGATCAAGTCGTATGCCACGCCACGATATCGCAAGAATACAGCTGGAGAAG	2600
QY	3125	GAGATGCACAAAGTGGACTTGCCCGATCAGACCCAAATCGAGATGAGGAANAATGCTCGAAC	3184
Db	2599	GAGATGCACAAAGTGGACTTGCCCGATCAGACCCAAATCGAGATGAGGAANAATGCTCGAAC	2540
QY	3185	CAAAAGGAGAGCAACTACATTCGATTAAGCGCGCCAAAGATGGACACAGACATGTTTCGTG	3244
Db	2539	CAAAAGGAGAGCAACTACATTCGATTAAGCGCGCCAAAGATGGACACAGACATGTTTCGTG	2480
QY	3245	AAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAACGCTGTGACCAAAATCGAT	3304
Db	2479	AAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAACGCTGTGACCAAAATCGAT	2420
QY	3305	ACCTCGAACCATTTGTATCGGATGAAAACCTTCGGGAAGCGGACGTTCTCAAGCGGAAT	3364
Db	2419	ACCTCGAACCATTTGTATCGGATGAAAACCTTCGGGAAGCGGACGTTCTCAAGCGGAAT	2360
QY	3365	CAGGTGGCACACTGAAGCCCGAGAGGATATCTTCGGGAAGCCGACATACTGGGTG	3424
Db	2359	CAGGTGGCACACTGAAGCCCGAGAGGATATCTTCGGGAAGCCGACATACTGGGTG	2300
QY	3425	GTGAAGTTGTACTACAGCTTCCAGGACAAGGATAATCTGTAATTTGTATGGACTACATA	3484
Db	2299	GTGAAGTTGTACTACAGCTTCCAGGACAAGGATAATCTGTAATTTGTATGGACTACATA	2240
QY	3485	C-----	3485
Db	2239	CCAGGTGAGTGGAACTTAAGCAGTCCAACGATGTATCCGGAGTCCAAATGTCATTGTG	2180
QY	3486	-----CAGGTGGTGATCTGATGTGCTGTGTCATCAAACTGGGCATTTTCGAGGAGG	3536
Db	2179	GATTTGCGACAGGTGTCATCTGATGTGCTGTGTCATCAAACTGGGCATTTTCGAGGAGG	2120
QY	3537	AACTGCCAGATTCACATCCGCGAGGTACCTGCGCGGTGGAGACGGTTCACAAATGG	3596
Db	2119	AACTGCCAGATTCACATCCGCGAGGTACCTGCGCGGTGGAGACGGTTCACAAATGG	2060

QY	3597	GCTTCATT-----	3604
Db	2059	GCTTCATTACAGGTAATTAGCAGTTGCTTCAGTTATCACAGTCACATTTACATATTGTT	2000
QY	3605	-----CACAGAGACATCAAGCCTGCACAACATCTCATCATGATAGGACGGACACA	3653
Db	1999	ATGGATTTCATCCACAGAGACATCAAGCTTGACAACATATCATCATGATAGGACGGACACA	1940
QY	3654	TAAAGCTCACCGACTTTGGCGCTGTGCACGGGATTCCGATGGAGCGCACAACTCGAAGTACT	3713
Db	1939	TAAAGCTCACCGACTTTGGCGCTGTGCACGGGATTCCGATGGAGCGCACAACTCGAAGTACT	1880
QY	3714	ACCAGGAGAAC-----	3724
Db	1879	ACCAGGAGACGGTGAGATGGAAACCCACGCTAACGTAGTGTGTTTACAGCTTTTAATAATAT	1820
QY	3725	-----GGCAATCACTCGCGCCAGGACTCGATGGAGGCCCTGGGAGGAATA	3768
Db	1819	CATTCCCTACTCTGATAGGCAATCACTCGCGCCAGGACTCGATGGAGGCCCTGGGAGGAATA	1760
QY	3769	CTCCGAGAACGGACCGGAAGCCCAACCGTGTGGAG-----	3802
Db	1759	CTCCGAGAACGGACCGGAAGCCCAACCGTGTGGAGAGGTGAGTTTCGTACACTTAAATTAAT	1700
QY	3803	-----AGCGGACGATCGCGATCACCAAA	3827
Db	1699	AGGATCGCGCGTTAATAAAGTATGCCGAACCTCGCAGCGGACGGATCGCGATCACCAAA	1640
QY	3828	GAGTCTGGCCCCACTCGCTGGTGGGCACCCGAACTACATAGCTCCGAGGTGCTGGAGA	3887
Db	1639	GAGTCTGGCCCCACTCGCTGGTGGGCACCCGAACTACATAGCTCCGAGGTGCTGGAGA	1580
QY	3888	GRAGTGGGTACGCGAGCTGTGGACTACTGAGCGTGGGCGTCATCTCTTATGAGATGC	3947
Db	1579	GAGTGGGTACGCGAGCTGTGGACTACTGAGCGTGGGCGTCATCTCTATGAGATGC	1520
QY	3948	TGTTGGGTACGCGCCCTTCTGGCCAAAGTCTCCGCTGGAACCAACAAA-----	3998
Db	1519	TGTTGGGTACGCGCCCTTCTGGCCAAAGTCTCCGCTGGAACCAACAAAAGTTTCGCT	1460
QY	3999	-----AGTCA	4004
Db	1459	CTTTGGATTCCCTTATAAGCCAAAGTGCAATTACAACATGACTTGTCTTATACAGGTCA	1400
QY	4005	TCAACTGGGAGAAACSCTGCAATTTCGCGCGCAGGCGGAGTTATCCGCGAGGCTACGG	4064
Db	1399	TCAACTGGGAGAAACCCCTGCAATTTCGCGCGCAGGCGGAGTTATCCGCGAGGCTACGG	1340
QY	4065	ACTTGATAGGAGGCTCTGTGGTGGCTGACACAGCGGCTGGGCAAGAGCGTGGACGAGG	4124
Db	1339	ACTTGATAGGAGGCTCTGTGGTGGCTGACACAGCGGCTGGGCAAGAGCGTGGACGAGG	1280
QY	4125	TCAAGAGCCACGACTTCTTCAAGGGCATCTGACTTTGGCGCATGTCGGGAACGACAAAGCCG	4184
Db	1279	TCAAGAGCCACGACTTCTTCAAGGGCATCTGACTTTGGCGCATGTCGGGAACGACAAAGCCG	1220
QY	4185	CTTACATACCGGAAATCAAGCACCCACGCGACACATCCAACTTTGATCCCGTGGATCCGG	4244
Db	1219	CCTATATACCGGAAATCAAGCACCCACGCGACACATCCAACTTTGATCCCGTGGATCCGG	1160
QY	4245	AGAAGCTGCGCTCGAATGACTCCACCATGACACGGCGATGATGTCGACCAGATGACC	4304
Db	1159	AGAAGCTGCGCTCGAATGACTCCACCATGACACGGCGATGATGTCGACCAGATGACC	1100
QY	4305	GCAYTTCACGGCTTTTTCGAAATTAACCTTCGCTCGCTTCTTCGACGACAAAGCGCCG	4364
Db	1099	GCACTTTCACGGCTTTTTCGAAATTAACCTTCGCTCGCTTCTTCGACGACAAAGCGCCG	1040
QY	4365	CGGATATGACGACGATCAGGCGCGGTTTACGTCCTCAATGATGCTCTCCATGTGCC	4424
Db	1039	CGGATATGACGACGATCAGGCGCGGTTTACGTCCTCAATGATGCTCTCCATGTGCC	980
QY	4425	AACACCAACACCCGCCGCCGAACTATTGTAGTCAAAATAGTCACAAAAAGGGGATAGAA	4484

Db 979 AACACCAACCCCGCCCGGAAATCATTTAGTCAAAATAGTACAAAAAGGGGATAGAA 920
QY 4485 ACCATTGAGTGGCTTGCTTAAAGGAAGCTGGCTATAGAAATGAACATCTATATA 4544
Db 919 ACCATTGAGTGGCTTGCTTAAAGGAAGCTGGCTATAGAAATGAACATCTATATA 860
QY 4545 CATTATATAAATTATAGGACAGTATAGAGCGGGAGCTAGTATATACATACAAATAA 4604
Db 859 CATTATATAAATTATAGGACAGTATAGAGCGGGAGCTAGTATATACATACAAATAA 800
QY 4605 TACATATATTGTAT 4663
Db 799 TACATATATTGTAT 740
QY 4664 AAAACGAGCGGAGTATAGATGAAA-CGAGAGGAGCGAGTCAGACCTTCGACCTTTAAC 4722
Db 739 AAAACGAGCGGAGTATAGATGAAAACGAGAGGAGCGAGTCAGACCTTCGACCTTTAAC 680
QY 4723 TGAACATAGTATATCTTGTGCACTACTACTCCACAAATATATATATATATATATAT 4782
Db 679 TGAACATAGTATATCTTGTGCACTACTACTCCACAAATATATATATATATATATAT 620
QY 4783 AGAATTCAAAAGGRCGAACCTGGAATCGAACCTTTCTGTGTCTCAAGCAAAAGCAAGC 4842
Db 619-AGAATTCAAAAGGAGCAACTGGAATCGAACCTTTCTGTGTCT-----CAAAGC 570
QY 4843 AAAGCAAAACAAACGCTTAACTAAAYGAGAGCGGAATTTACCAACACACCTTCACTCC 4902
Db 569 AAAGCAAAACAAACGCTTAACTAAACGAGAGCGGAATTTACCAACACACCTTCACTCC 510
QY 4903 TCTCTTTTCCACCTCCGATCGGTGGCGGATTCGAACCTCAGCAGCTGGTTGCATCCG 4962
Db 509 TCTCTTTTCCACCTCCGATCGGTGGCGGATTCGAACCTCAGCAGCTGGTTGCATCCG 450
QY 4963 GCATCCCATTTACTTCCCATTCAGAAATGAGATTCGAGGTGTGGATGGAGAGCAAC 5022
Db 449 GCATCCCATTTACTTCCCATTCAGAAATGAGATTCGAGGTGTGGATGGAGAGCAAC 390
QY 5023 GGAGACCAAAAGTCGACGCGCATATAGCGGCTTATAGCCTATCTTAATCTA 5082
Db 389 GGAGACCAAAAGTCGACGCGCATATAGCGGCTTATAGCCTATCTTAATCTA 330
QY 5083 AACTGGGAGAACAGGACCTATGTA-----TGCTCTGTCTCCATTCGTCTATC 5131
Db 329 AACTGGGAGAACAGGACCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 270
QY 5132 ACTGCTCTCAWCTGWTAGACCCGCCACCCGCC-----CTCCCATCCAAAGAACAA 5187
Db 269 ACTGCTCTCAWCTGWTAGACCCGCCACCCGCC-----CTCCCATCCAAAGAACAA 210
QY 5188 ACTTAGACCTAGCTATGTCAAAAGCTAGCAATGTTAGACCAACTTGTGAATGCCAAAT 5247
Db 209 ACTTAGACCTAGCTATGTCAAAAGCTAGCAATGTTAGACCAACTTGTGAATGCCAAAT 150
QY 5248 GAAATGTTTATAGCCCAACGAGGAAAACGCGGGGAAATTCACACATTAATCTCTGATAGC 5307
Db 149 GAAATGTTTATAGCCCAACGAGGAAAACGCGGGGAAATTCACACATTAATCTCTGATAGC 90
QY 5308 AAACGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAA 5367
Db 89 AAACGAAAAGAAAAG-----AAAAAAGAAAAGAAAAGAAAAGAAAAGAAAAGAAAAG 35
QY 5368 TTAATGTAATATTGTAAGAACACGTTAATTGTA 5401
Db 34 TTAATGTAATATTGTAAGAACACGTTAATTGTA 1

RESULT 4
ID ABL03169
ABL03169 standard; cDNA; 3319 BP.
XX
AC ABL03169;

XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 3989.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
PN 27-SEP-2001.
PD 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
XX P-PSDB; ABB59066.
DR New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions.
PS Claim 1; SEQ ID NO 3989; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 3319 BP; 860 A; 1032 C; 933 G; 494 T; 0 other;
SQ
Query Match 56.7%; Score 3242.2; DB 23; Length 3319;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3276; Conservative 16; Mismatches 9; Indels 18; Gaps 2;
QY 1103 ATGCATCCAGCGCGGAAAAGGGGGGCGGTCGCCCAATGATAAATACACGCGGAGCC 1162
Db 1 ATGCATCCAGCGCGGAAAAGGGGGGCGGTCGCCCAATGATAAATACACGCGGAGCC 60
QY 1163 CTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGTACAAAATACCATAGGAATAAT 1222
Db 61 CTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGTACAAAATACCATAGGAATAAT 120
QY 1223 CAGAAATACACCTCTGCGATACAGCGGACCAACGAGCAGCATGACCTTACTCTCT 1282
Db 121 CAGAAATACACCTCTGCGATACAGCGGACCAACGAGCAGCATGACCTTACTCTCT 180
QY 1283 GACTATCACACCGCAGCAGCATGAGCGCCACCTCCGCCCTCTCTCTCTCTCTCTCT 1342
Db 181 GACTATCACACCGCAGCAGCATGAGCGCCACCTCCGCCCTCTCTCTCTCTCTCTCT 240
QY 1343 GTGGTATACCGCGCGCGCCGATTTAGGTGACGCCGCGGCTCCATATCCGTA 1402
Db 241 GTGGTATACCGCGCGCGCCGATTTAGGTGACGCCGCGGCTCCATATCCGTA 300
QY 1403 TCCGGTGTGGCGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1462
Db 301 TCCGGTGTGGCGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 360

Qy	1463	GCCTTAATGCCAAACAAATGATCGGAAGCGAGCATCGAAGCGGACACGCGGAGCACT	1522
Db	361	GCCTTAATGCCAAACAAATGATCGGAAGCGGAGCATCGAAGCGGACACGCGGAGCACT	420
Qy	1523	CACTACCTCGGCTCGAGTCGGCTCTGGACTCGGAGCCGGTAGCTCCCGATCGGACAGC	1582
Db	421	CACTACCTCGGCTCGAGTCGGCTCTGGACTCGGAGCCGGTAGCTCCCGATCGGACAGC	480
Qy	1583	CCCATTCGCACACACACCCACGAGCTCGAGGACGGTGGTAAATCCAGGTGGAAT	1642
Db	481	CCCATTCGCACACACACCCACGAGCTCGAGGACGGTGGTAAATCCAGGTGGAAT	540
Qy	1643	GGTGGATTTTCCTCGTCGCAAGCGTTTCAGTGAGGTGGTCCACCGCGCGCGCCCA	1702
Db	541	GGTGGATTTTCCTCGTCGCAAGCGTTTCAGTGAGGTGGTCCACCGCGCGCGCCCA	600
Qy	1703	CGCAATCCCAACCCGCTSCAGCGCGGCGCCGCCCGCCAGTCGCGGCCACACGACCGAG	1762
Db	601	CGCAATCCCAACCCGCTSCAGCGCGGCGCCGCCCGCCAGTCGCGGCCACACGACCGAG	660
Qy	1763	GCCTAGCTGAAGCGGGATCACCGGCCCTGAACAACCGCCCGCGGGATAGCGCCACCC	1822
Db	661	GCCTAGCTGAAGCGGGATCACCGGCCCTGAACAACCGCCCGCGGGATAGCGCCACCC	720
Qy	1823	ACTCAGCGMGGAACCTCACCTGTAATACCCARAACGGGCTGAAGAACCCGCGACGACGAG	1882
Db	721	ACTCAGCGMGGAACCTCACCTGTAATACCCARAACGGGCTGAAGAACCCGCGACGACGAG	780
Qy	1883	TTGACGACGAGCTGAAGTCCCTGAACCTATACCCAGCGGAGGAGCTGGAGCAGTGTG	1942
Db	781	TTGACGACGAGCTGAAGTCCCTGAACCTATACCCAGCGGAGGAGCTGGAGCAGTGTG	840
Qy	1943	GAGCCACGCGCCCTACCTAAATTCAGGCGGAGCGGAGGAGCAGCACCGCGCGCCCA	2002
Db	841	GAGCCACGCGCCCTACCTAAATTCAGGCGGAGCGGAGGAGCAGCACCGCGCGCCCA	900
Qy	2003	CCACCCAGTTACAGGGCTCCATGCACTCGGGGAGTGGGCCACACAATCCCAACATCG	2062
Db	901	CCACCCAGTTACAGGGCTCCATGCACTCGGGGAGTGGGCCACACAATCCCAACATCG	960
Qy	2063	GACTACAGGAATCCCGAGCAGTGGATATCTCGGCCACCTCGCGGGGCTCGCCGAGC	2122
Db	961	GACTACAGGAATCCCGAGCAGTGGATATCTCGGCCACCTCGCGGGGCTCGCCGAGC	1020
Qy	2123	CCCAATACTGTGCTGCTCGCGCGCGCTGCGGAAGCACAACACAGTCTACCA	2182
Db	1021	CCCAATACTGTGCTGCTCGCGCGCGCTGCGGAAGCACAACACAGTCTACCA	1080
Qy	2183	GCCAGGAGTCAGACCGCATCATCAGAGTGTGAAGACAGCGAGTCCAAAAGGCC	2242
Db	1081	GCCAGGAGTCAGACCGCATCATCAGAGTGTGAAGACAGCGAGTCCAAAAGGCC	1140
Qy	2243	GTGCTGCAACACAGAGTGGCGGCGCAATCGCCATCGAGTGTGCTCGGCGAGCAATTCACCA	2302
Db	1141	GTGCTGCAACACAGAGTGGCGGCGCAATCGCCATCGAGTGTGCTCGGCGAGCAATTCACCA	1200
Qy	2303	GTCCAGTGTGCGGCTTCCACCTTACCTCAGAGTCCGCGGAGTGGTCAGAG	2362
Db	1201	GTCCAGTGTGCGGCTTCCACCTTACCTCAGAGTCCGCGGAGTGGTCAGAG	1260
Qy	2363	CAGCAACAGGACAGCGCGCGCCACCCAGGACGAGCATCAGCACAGCAATCCAAACCA	2422
Db	1261	CAGCAACAGGACAGCGCGCGCCACCCAGGACGAGCATCAGCACAGCAATCCAAACCA	1320
Qy	2423	CCACGCCMACCACACCGCCCTTGTGGTCTGAAYACGACCCCAATTGCTTGAGCCA	2482
Db	1321	CCACGCCMACCACACCGCCCTTGTGGTCTGAATAGCAAGCCCAATTGCTTGAGCCA	1380
Qy	2483	CGGCTCTATGCCAAGAGCATGAGGCCAAGGCGCGCGGCTGGTGAAGACGACGACAA	2539
Db	1381	CGGCTCTATGCCAAGAGCATGAGGCCAAGGCGCGCGGCTGGTGAAGACGACGACAA	1440

Db 2521 CACAGACATCAAGCTGCAACATATCTATCATGATAGGACGACACATAAAGCTCACC 2580
 QY 3665 GACTTTGGCTGTGCACGGATTCCGATGGACGACACAACTCGAAGTACTACCGAGAGAAC 3724
 Db 2581 GACTTTGGCTGTGCACGGATTCCGATGGACGACACAACTCGAAGTACTACCGAGAGAAC 2640
 QY 3725 GGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGGAGGAATACTCCGAGAACGGACCG 3784
 Db 2641 GGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGGAGGAATACTCCGAGAACGGACCG 2700
 QY 3785 AAGCCACCTGCTGGAGAGCGGAGGATGGCGATCACCAGAGTCTTGCCCATCTCG 3844
 Db 2701 AAGCCACCTGCTGGAGAGCGGAGGATGGCGATCACCAGAGTCTTGCCCATCTCG 2760
 QY 3845 CTGGTGGGACCCCGCACTACATAGCTCCGAGGTCTGGAGAGGAGTGGGTACACCCAG 3904
 Db 2761 CTGGTGGGACCCCGCACTACATAGCTCCGAGGTCTGGAGAGGAGTGGGTACACCCAG 2820
 QY 3905 CTGTGGCACTACTGGAGCGTGGGCGTCTCTTAYAGATGCTGTGGGTTCAGCCGCC 3964
 Db 2821 CTGTGGCACTACTGGAGCGTGGGCGTCTCTTAYAGATGCTGTGGGTTCAGCCGCC 2880
 QY 3965 TTTCTGCCAACAGCTCGCTGGAAACGCAACAAAGTCTCAACTGGGAGAAACGCTG 4024
 Db 2881 TTTCTGCCAACAGCTCGCTGGAAACGCAACAAAGTCTCAACTGGGAGAAACGCTG 2940
 QY 4025 CATATTCGCGCGGAGCGGAGTATCCGCGAGGCTACGACTGTGAAGGAGGCTCTGT 4084
 Db 2941 CATATTCGCGCGGAGCGGAGTATCCGCGAGGCTACGACTGTGAAGGAGGCTCTGT 3000
 QY 4085 GCCTGGCTGACAGCGGCTGGCAAGAGGCTGGACAGGTCAAGAGCCACGACTTCTTC 4144
 Db 3001 GCCTGGCTGACAGCGGCTGGCAAGAGGCTGGACAGGTCAAGAGCCACGACTTCTTC 3060
 QY 4145 AAGGCATGCACTTTGGGACATGCGGAGAGCAAGAACGCGCTTACATACCGGAAATCAAG 4204
 Db 3061 AAGGCATGCACTTTGGGACATGCGGAGAGCAAGAACGCGCTTACATACCGGAAATCAAG 3120
 QY 4205 CACCCACGACACATCCAACTTTGATCCGCTGATCCGAGAGGCTCGCTCAATGAC 4264
 Db 3121 CACCCACGACACATCCAACTTTGATCCGCTGATCCGAGAGGCTCGCTCAATGAC 3180
 QY 4265 TCCACCATGACAGCGCGGATGATGTCAGCAGAAATGACCGCCTTCCACGGCTTTTC 4324
 Db 3181 TCCACCATGACAGCGCGGATGATGTCAGCAGAAATGACCGCCTTCCACGGCTTTTC 3240
 QY 4325 GAATTTACCTTCGCTGCTTCTTCGAGCAAGCAGCGCGCGGATATGACGGAGCATGAC 4384
 Db 3241 GAATTTACCTTCGCTGCTTCTTCGAGCAAGCAGCGCGCGGATATGACGGAGCATGAC 3300
 QY 4385 GCGCGGTTTACGCTCTGAA 4403
 Db 3301 GCGCGGTTTACGCTCTGAA 3319

RESULT 5

AAAT42120
 ID AAAT42120 standard; cDNA; 3155 BP.
 XX
 AC AAAT42120;
 XX
 DT
 XX
 XX 31-JAN-1997 (first entry)
 XX M-lats2 gene encoding large tumour suppressor.
 DE
 XX Mouse; m-lats2 gene; large tumour suppressor; fetal brain;
 KW protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX
 OS Mus musculus.

XX FH Key Location/Qualifiers
 FT CDS 1..2943
 FT /*tag= a
 FT /product= m-lats2 protein
 XX
 PN WO9630402-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 26-MAR-1996; 96WO-US04101.
 XX
 PR 27-MAR-1995; 95US-0411111.
 XX
 PA (UYUA) UNIV YALE.
 XX
 PI Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX
 DR WPI; 1996-455275/45.
 DR P-PSDB; AAW06287.
 XX
 PT New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Claim 8; Page 133-137; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats2
 CC protein, and has been isolated from a newborn mouse brain phase
 CC lambda-ZAP cDNA library using a 2.2-kb DNA probe from the Drosophila
 CC lats gene (AAT42117). A homologous mouse sequence has also been
 CC isolated (m-lats, AAT42119). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;
 Query Match 9.9%; Score 567.8; DB 17; Length 3155;
 Best Local Similarity 64.3%; Pred. No. 6.3e-110;
 Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;
 QY 2869 GGGCGGAATAGCTCGGGGGGAGCAACGGATCCCGGACACCGCCTCTCTCGTGCAC 2928
 Db 1329 GGGCACTGATCTAGACGGGAGTGACAAGAGCCAAAGGTGCGAAGGAGAGCAAGCTGG 1388
 QY 2929 CAGCTGCAAGAAGATCAAGCAGCGCTCGCCATCCCGGAGGCGGAGAGATCTCCAAGGA 2988
 Db 1389 CAGAGACAAAAGAGCAGATTCAGACCTCCCGGTGCTCTCGCAAGAAATAGC----- 1440
 QY 2989 GAAGGAGGAGGAGCGCAAGGAGTTCCGATCAGGAGTCTCGCGCAAGCCTTCAAGTT 3048
 Db 1441 -AGAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATCGCTTCAATT 1499
 QY 3049 CTTTCATGGAGCAGCATATAGAACGTTGATCAAGTCTGATCGCAGCGCAGTATCGCAA 3108
 Db 1500 CTTTCATGGAGCAACACGCTGGAGAATGTATCAAAACCTACGAGCAAGAGGTGACGCCGAG 1559
 QY 3109 GAATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATCGAGAT 3168
 Db 1560 GCTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGGCGCAGCAGGAGCAGAT 1619
 QY 3169 GAGGAAATGCTGAACCAAGAGGAGAGCAACTACATTCGATTGAAGCGCCGAAGATGGA 3228
 Db 1620 GAGGAGATCTCTTACCGAGAGGAGTCTAATCAACCGGCTGAAGAGGCGCAAGATGGA 1679
 QY 3229 CAAGAGCATGTTCTGTTAACTGAAGCCCAATTTGGAGTGGGTGCAATTTTGGCGAGGTAAAGCT 3288

Db 1680 CAAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCCTTTGGGAAAGTGTGCCT 1739
QY 3289 GGTGAGCAAAATCGATACCTCGAACCATTTGTATGGATGAAACCCCTGCGGAAAGCGGA 3348
Db 1740 CGCTTGTGAAGCTGGA---CACTCAGCGCTGTGTAGCCATGAAGACTCTCAGGAAGGA 1796
QY 3349 CGTCTCAAGCGGAATCAGGTGCGACACGCTGAAGCGCGAGAGGATATCTCCGCGAAGC 3408
Db 1797 TGTCTTGAACCGGAATCAAGTGGCCATGTCAAGGCTGAGAGGACATCTGCTGCTGAAGC 1856
QY 3409 CGAATAAATCGTGGTGGTGAAGTGTACTACAGTCTCCAGGCAAGGATATCTGTACTT 3468
Db 1857 AGACATGAGTGGTGGTGAAGTGTACTACTCTTCCAGGCAAGGACAGCGCTGTACTT 1916
QY 3469 TGTGATGGACTACATACCAAGGTGGTGTATGTGTGCTGCTCATCAAACTGGGCAATTT 3528
Db 1917 TGTGATGGACTACATACCAAGCGGGGATATGATGAGCTGCTGATCAGGATGAGTCTT 1976
QY 3529 CGAGGAGAACTGGCCAGATTTACATCGCGAGGTCTACCTGGCCGCTGAGACAGCTTCA 3588
Db 1977 CCTGAGCACCTGGCCGCTTCTACATTCAGAGTTGACCCCTGGCCATTTGAAAGTGTCCA 2036
QY 3589 CAAATGGGCTTCATTACAGAGACATCAAGCTGACACATCTCATCATGATAGGACGG 3648
Db 2037 CAAGATGGGCTTTATCCACCGGGACATCAAGCTGACACATCTCATCATGATAGG 2096
QY 3649 ACATATAAGCTACCGACTTTGGCTGTGCACGGGATTCGATGGACGACACAACTCGAA 3708
Db 2097 TCATATAAGCTGACAGATTTTGGCTGTGCACGGGATTCAGTGGACTCACAATTCGAA 2156
QY 3709 GTACTACAGGAGAACCGCAATCACTCGCGCAGGACTCGATGGAGCCCTGGGAGAAAT- 3767
Db 2157 GTACTACC---AGAAAGGAAACCATATGACAGGACAGCATGGAGCCCGGTGACTCTG- 2213
QY 3768 -----ACTCCGGAAGACCGGACCGCCAGTGTGGAGAGCGGAGAT 3813
Db 2214 GGACGATGTTTCCACTGTGCTGTGGACAGAGTTAAGACCTTGGACAGAGGCGGA 2273
QY 3814 GCGGATACCAAGAGTCTTGGCCACTGCTGTGGGACCCCGGAACTACATAGCTCC 3873
Db 2274 GAAGCAGCACAGAGTGTGCGACATTTCTTGTGGGACACCAAAATACATCGCTCC 2333
QY 3874 CGAGTGTCTGGAGAGAGTGTGCTGACGAGCTGTGCGGACTACTGGAGCGTGGGCTCAT 3933
Db 2334 GGAGTGTCTTCCGAAAGGAGTACGAGCTGTGTGACTGGTGGAGCGTGGTGTGAT 2393
QY 3934 CCTTAYAGAGTCTGGTGGTTCAGCGCCCTTCTTGGCCACAGTCCGCTGGAACGCA 3993
Db 2394 TCTCTTTGAGATGCTGTGGCAGCGCCCTTCTTGGCCCGCCACCCACAGACAGCA 2453
QY 3994 ACAAGGTATCACTGGGAGAAACSTGATATTCGCGCGAGCGGAGTATTCGG 4053
Db 2454 GCTGAAGTGTATCACTGGGAGACAGCTGTATATCCCTAGCGAGTGTAGGCTCAGCGC 2513
QY 4054 CGAGCTACGAGTGTATGAAGAGAGCTGTGCTGCTGCTGACAGCGCTGGGCA---A 4110
Db 2514 TGAGCCGAGACCTCATACAGAGCTGTGTGCGCGCTGACTCCCGCTGGGAGGGA 2573
QY 4111 GAGCGTGGACAGGTCAAGAGCAGGCTTCTTCAAGGCGCATCGACTTTC---GGACAT 4167
Db 2574 TGGGCGAGATGACCTCAAGGCACACCCGCTTCTCAACACCATCGACTTTTCCGCTGAT 2633
QY 4168 GGGAGCAGAAAGCGCCCTCATACCGGAATCAAGCACCCACCGGACACATCCAACTT 4227
Db 2634 CGGAAGAGGCTGACCCCTAGTCCCGCCACCATCAGCCACCCCATGGACACTCCAACTT 2693
QY 4228 TGATCCGCTGATCGGAGAGCTGCGCTCGAATGACTCCACCATGACAGCGGCGGATGA 4287
Db 2694 TGACCCGCTGATGAAGAAAGCCCTGCGCAGAGGCGGAGAGCGGCAAGGCGCTG 2753
QY 4288 TGTCCACAGAAATGACCCACATTTTC-----CACGGCTTTTTCGAATTTACCTT 4335
Db 2754 GGACAGCTGCGCTCCCGCCAGCAGCAAGCATCCAGAGGACCGCTTCTATGATGATCCTT 2813

QY 4336 CCCTCGCTTCTTCGACGACAA 4356
Db 2814 CCAGAGTCTTCGATGACAA 2834

RESULT 6

AAZ51507
ID AAZ51507 standard; DNA; 3155 BP.

XX AAZ51507;
AC AAZ51507;

XX 21-JUN-2000 (first entry)
DT XX

XX Mouse Lats2 (large tumour suppressor) DNA.

XX Mouse; Lats2; large tumour suppressor; cytostatic; vulnary;
KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
KW treatment; prevention; screening; cancer; skin; ovarian tumour;
KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
KW LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX OS Mus musculus.

XX Key Location/Qualifiers
FH 1. 2943
CDS

FT /tag= a
FT /product= "Lats2 protein"

FT misc_feature
FT 970..1920
FT /tag= b

FT /note= "This region is erroneously repeated in the
mouse Lats2 DNA sequence shown in figure 14"

FT misc_feature
FT 1921..2960
FT /tag= c

FT /note= "This region is missing in the mouse Lats2
DNA sequence shown in figure 14"

FT WO200010602-A1.

XX 02-MAR-2000.

XX 18-AUG-1999; 99WO-US19068.

XX 18-AUG-1998; 98US-0096996.

XX 18-AUG-1998; 98US-0096997.

XX (UYUA) UNIV YALE.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
PI Trenchalk GS, Stewart RA;

XX WPI: 2000-246496/21.
P-PSDB: AAY70392.

XX Use of lats proteins, complexes of lats and cdc2 for treating cancer
that is refractory to treatment by standard chemotherapy and radiation
therapy, and disorders associated with aberrant levels of cdc2 activity

XX Claim 44; Pages 112-117; 134pp; English.

XX The present sequence is a DNA encoding mouse Lats2 (large tumour
suppressor) protein which is a cell overproliferation inhibitor and a
negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
XX The present sequence is useful for treating cancer that is refractory
to standard chemotherapy or radiation therapy such as hyperplasia,
metaplasia, or dysplasia, and disorders associated with aberrant
levels of cdc2 activity. Conditions treated by promoting cdc2 function
include degenerative disorders, growth deficiencies, hypoproliferative
disorders, physical trauma, lesions, and wounds. An animal model
preferably a mouse, in which a lats gene has been disrupted by homologous

CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX

Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;
 Query Match 9.9%; Score 567.8; DB 21; Length 3155;
 Best Local Similarity 64.3%; Pred. No. 6.3e-110;
 Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCCGGAATAGTCGGGCGGAGCAAGGATCCACCGGACACCGCTCTCTCGTCGAC 2928
 DB 1329 GGGCACTGATCTAGACGGGAGTGACAAGAGCCACAAAGGTCGAAGGGAGAGCAAAAGCTGG 1388
 QY 2929 CAGCTGCAAGAGATCAAGCAGCGCTCCGCCATCCCGGAGCCGCAAGATCTCAAGGA 2988
 DB 1389 CAGAGCAAAAGAGATTCAGACCTCCCGGTGCTGTCCGCAAGATAGC----- 1440
 QY 2989 GAAGGAGGAGGAGCGCAAGGAGTTCCGCATCAGGCGAGTACTCCCGCAAGCTTCAAGTT 3048
 DB 1441 -AGAGATGAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTATGCTTCAAAAT 1499
 QY 3049 CTTGATGAGCAGCAGATAGAGAACTGATCAAGTCGATCCCGCAGCGCAGATATCGCAA 3108
 DB 1500 CTTGATGAGCAGCAGATAGAGAACTGATCAAGTCGATCCCGCAGCGCAGATATCGCAA 3108
 QY 3109 GAATCAGCTGAGAGGAGATGACAAAGTGGGACTCCCGATCAGACCCAAATCGAGAT 3168
 DB 1560 GCTACAGCTGAGAGGAGGAGTGGGCTGCTGAGGCGGAGGAGGAGCAGAT 1619
 QY 3169 GAGGAAATGCTGAAACCAAGGAGGAGCACTACATTCGATTGAAGCGCGCAAGATGGA 3228
 DB 1620 GAGGAGATCTCTACAGAGAGAGTCTAATACACCGCTGAGAGGCGCAAGATGGA 1679
 QY 3229 CAAGAGCATGTTGCTCAAACTGAAGCCGATGGAGTGGGTCATTTGGCGAGGTAAACGCT 3288
 DB 1680 CAAGTCCATGTTGTTGAAAATCAAGACTCTAGGCACTGCTGCTTTGGGGAAGTGTGCT 1739
 QY 3289 GGTGAGCAAAATCGATCTCGAACCATTGTTATCGGATGAAACCTTCGGGAAAGCGGA 3348
 DB 1740 CGTTGTAAGCTGGA---CACTCAGGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGGA 1796
 QY 3349 CGTTCTCAAGCGGAATCAGGTGGACACGCTGAAGCGCGGAGGAGATATCTTCGCGGAAGC 3408
 DB 1797 TGTCTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGAGATCTTCGGTGAAGC 1856
 QY 3409 CGACAATACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGGATATCTGTACTT 3468
 DB 1857 AGACAATGAGTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGGAGAGCTGTACTT 1916
 QY 3469 TGTGATGACTACATACAGGTGGTGTGATCTGATGCTGCTCATCAACTGGGCGATTTT 3528
 DB 1917 TGTGATGACTACATACAGGTGGTGGTGTGATGATGAGCTGTGTGATCAGGATGGAGGTCT 1976
 QY 3529 CGAGGAGCAACTCGCCGATTTCTACATCGCGAGGTCACTCGCGCGGTGGAGAGGTTCA 3588
 DB 1977 CCCTGAGCACTGGCCGCTTCTACATTCGAGATTGACCTTCCAGGAGTGAAGTGTCA 2036
 QY 3589 CAAATGGCTTCAATTCACAGAGACATCAAGCCTGACAACTATCTCATCGATAGGAGCGG 3648
 DB 2037 CAAATGGCTTCAATTCACAGAGACATCAAGCCTGACAACTATCTCATCGATAGGAGCGG 2096
 QY 3649 ACACATAAGCTCAGGCTTGGCTGTGACGGGATTCGATGGAGCGCAACATCTCAA 3708
 DB 2097 TCATATTAAAGTGACAGATTTGGCCCTCTGCACCTGGATTTCAGGTGGAGTCACAATTCCAA 2156
 QY 3709 GTACTTACAGGAGCAAGCAATCACTCGCGCAGGAGTCTGATGGAGCGCTGGGAGGAT- 3767
 DB 2157 GTACTACG---AGAAAGGAGCAACATGAGACAGAGCATGGAGCCCGGTGACCTCTG 2213
 QY 3768 -----ACTCCGAGAACGGACCGAAGCCACCGTGTCTGGAGAGCGGACGGAT 3813

DB 2214 GGACGATGTTTCCAACTGTGCGTGTGGAGACAGGTTAAAGACCCCTGGAGCAGGCGCA 2273
 QY 3814 GGGCCATCACCACAAAGAGTCTGGCCACTCGCTGTGGGACCCCGAAGTACATAGCTCC 3873
 DB 2274 GAAGCAGCACCAGAGGTGCTTGGCAGATCTCTTGTGGGACACCAAAATTACATCGCTCC 2333
 QY 3874 CGAGTGTCTGGAGAGAGTGGGTACACGAGCTGTGCGACTACTGAGGCTGGGCTCAT 3933
 DB 2334 GGAGTGTCTTCCGCAAGAGGTACACGAGCTGTGACTGTGGAGGCTCGGTGTGAT 2393
 QY 3934 CTTTAYGAGATGCTGGGTGAGCCGCTTTCTGGCCAAACAGTCCGCTGGAAACGCA 3993
 DB 2394 TCTCTTTGAGATGCTGGTGGCAGCGGCTTTCTTGGCCCCACCCACACAGACGCA 2453
 QY 3994 ACAAAAGTCTCAACTGGGAGAAAACSTGCATATTCGCGCGCAGGCGGATTCCTCCG 4053
 DB 2454 GCTGAAGTGTATCAACTGGGAGAGCAGCTGCATATCCCTACGCAAGTGTGAGCTCAGCGC 2513
 QY 4054 CGAGCTACGGAATTAAGAGGAGCTGTGCGTGGCTGACAAAGCGCTGGGCA---A 4110
 DB 2514 TGAGCCCGGAGACCTCATCAGCAAGCTGTGCTGGCGGCTGACTGCGGCTGGGAGGGA 2573
 QY 4111 GAGCTGGAAGAGTCAAGAGCCAGCACTTTCTTCAAGGCGCATCGACTTTGC---GGACAT 4167
 DB 2574 TGGGCGAGATGACCTCAAGGACACCCGTTCTTCAACACCATCGACTTTTCCCGTGACAT 2633
 QY 4168 GGGGAAGCAGAAAGCGCTTACATACCGGAATCAAGACCCAGGAGGAGACATCAACTT 4227
 DB 2634 CGAAAGAGGCTGACCCCTAGTCCCAACCATAGCCACCCATGGACACCTCCAATTT 2693
 QY 4228 TGATCCCGTGGATCCGAGAGAGCTGCGTTCGAATCACTCCACCATGAGCAGGCGGATGA 4287
 DB 2694 TGACCCGCTGGATGAGAAAGCCCTGGCAGCCAGGAGAGAGGCGCAAGGCGCTG 2753
 QY 4288 TGTGACACAGATGACCGACACTTTC-----CACGCGCTTTTTCGAATTTTACCTT 4335
 DB 2754 GGACACGCTGGCTTCCCGCAGCAGCAAGCATCCAGCAGCAGCGCTTCTATGAGTTTACCTT 2813
 QY 4336 CGCTGCTTCTTCGAGGACAA 4356
 DB 2814 CGCAGGTTCTTCGATGACAA 2834

RESULT 7
 AAX87397
 ID AAX87397 standard; cdna; 5276 BP.

XX AAX87397;

AC AC
 XX 08-OCT-1999 (first entry)

XX Human WART2 cdna.

XX WART2; hWART2; WART orthologue; human; signal transduction;
 KW protein kinase; cancer; tumour; diagnosis; therapy; ss..

XX Homo sapiens.

XX Key Location/Qualifiers

FT 5'UTR 1..374

FT CDS /*tag= a

FT 3'UTR 375..3641

FT /*tag= b

FT /*tag= c

PN WO9937787-A2.

XX 29-JUL-1999.

XX 20-JAN-1999; 99WO-US01145.

PR 21-JAN-1998; 98US-0072023.
 XX (SUG-) SUGEN INC.
 XX Flanagan P, Plowman GD;
 XX WPI; 1999-458698/38.
 DR P-PSDB; AAY06527.
 XX
 XX New nucleic acid encoding human orthologs of Drosophila WART
 PT proteins, used to identify specific modulators for treating cancer
 PT or for diagnosis
 XX
 PS Claim 1; Page 122-123; 137pp; English.
 XX
 CC This is the nucleotide sequence of a cDNA clone coding for a human
 CC orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor
 CC serine/threonine kinase WART. hWART2 cDNA was isolated from a
 CC human bone marrow cDNA library using a PCR fragment of hWART1 as
 CC probe. hWART2 is consistently expressed in human tumour cells
 CC lines, except for most of the colon cancer lines examined. This
 CC overexpression in tumour cells versus normal tissues may provide a
 CC target for oncology drug development. Nucleic acids encoding
 CC full-length hWART2 and hWART2 polypeptides lacking one or more of
 CC amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518,
 CC and 974-1048, or lacking one or more of the N-terminal domain,
 CC catalytic domain, or C-terminal domain are claimed, as well as
 CC hWART1 sequences (AA87396), hWART2 and hWART1 polypeptides,
 CC antibodies, a method for identifying modulators of hWART function,
 CC and use of such modulator compounds to treat an abnormal condition
 CC involving hWART signal transduction, especially cancer. Probes for
 CC detection of hWART nucleic acids are also claimed.
 XX
 SQ Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;

Db 2498 TCAGGTGGCCCAACGCTCAAGGCCGAGAGGGACATCTCTGGCCGAGGACACATAGTGGGT 2557
 QY 3424 GGTGAAGTTGTACTACAGCTTCCAGGACAAGGATAAATCTGTACTTTTGTGTGGACTACAT 3483
 Db 2558 GGTCAAACTCTACTCTCTCTTCCAGGACAAGGACAGCTGTACTTTGTGTGGACTACAT 2617
 QY 3484 ACCAGTGGTGTATCTGATCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACCTGC 3543
 Db 2618 CCCTGGTGGGACATGATGAGCCCTGCTGATCCGGATGGAGGTCTTCCCTGAGCACCTGGC 2677
 QY 3544 CAGATTCTACATCCCGAGGTCACTCCGCGGTGGACAGCGTTCACAAAATGGCTTCAT 3603
 Db 2678 CCGGTTCTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCACAGATGGCTTCAT 2737
 QY 3604 TCACAGAGACATCAAGCTGACAACTACTCATCATGATAGGAGCGACACATAAAGCTCAC 3663
 Db 2738 CCACCGAGACATCAAGCTGATACATTTGATAGATCTGGATGTCACATTAACCTCAC 2797
 QY 3664 CGACTTTGGCTGTGACGCGGATTTCCGATGACGACCACTCGAAGTACTACACAGAGA- 3722
 Db 2798 AGATTTCGGCTCTGCACTGGTTTCAGTGCATCAAAATTTCCAAATATTACACAGAAAG 2857
 QY 3723 -----ACGGCAATCACTCGCGCCAGGACTCATGATGAGCCCTGGGAGGAATACTCCGA--- 3774
 Db 2858 GAGCCATGTGACAGACAGACATGAGCCAGGACCTCTGGGATGATGTCTAACTG 2917
 QY 3775 ---GAACGGACCGAAGCCACCGTGTGGAGAGCGGAGTGGCGATCACCACAAAGT 3831
 Db 2918 TCGGTGTGGGACAGGCTGAAGACCTTAGACAGAGGCGGGAAGCAGCACCAGAGGTG 2977
 QY 3832 CTTGGCCACTCGCTGTGGGACCCCGCAACTACATAGCTCCCGAGGTGCTGGAGAGAG 3891
 Db 2978 CTTGGCACATTCATCTGTGGGACTTCCAAACTACATCGACCCGAGGTGCTCTCCCGCAA 3037
 QY 3892 TGGGTACACGAGCTGTGCACTACTGAGGCGTGGGCGTATCTCTYAGAGATGCTGGT 3951
 Db 3038 AGGTACATCAACTGTGACTGTGGAGTGTGGAGTGTATCTCTTCGAGATGCTGGT 3097
 QY 3952 GGTGACGCGCCCTTCTGCGCAACAGTCCGTGGAACGCAACAAAGGTTCATCAACTG 4011
 Db 3098 GGGCAGCCGCGCTTTTGGCACTTCTCCACAGAACCCAGCTGAAGGTGATCAACTG 3157
 QY 4012 GGAGAAACCTGCATATTCGCGGAGCGGAGTATCCCGGAGGCTACGGACTTGTAT 4071
 Db 3158 GGAGAACACGCTCCACATTCAGCCAGGTGAAGCTGAGCCCTGAGCCGAGGACCTCAT 3217
 QY 4072 AAGGAGCTCTGTGCTGCTGACAGAGCGCTGGGCAAGCGCTGCAAGG---TCAA 4128
 Db 3218 CACCAAGCTGTGCTGCTCCGCGACACCCGCTGGGCGGAATGGGCGCGATGACCTGAA 3277
 QY 4129 GAGCCAGCTTCTTCAAGGCGATCGACTTTCG---GGACATCGGAGGAGCAAGCGCC 4185
 Db 3278 GGGCCACCCCTTCTTCAAGGCGATGACTTCTCCAGTGACATCGGAGGAGCAGCGCC 3337
 QY 4186 CTACATACCGGAAATCAAGCACCCACGACACATCACTTTGATCCCGTGGATCCGGA 4245
 Db 3338 CTAGTTCCCAACCATCAGCCACCCATGGACACCTCGAATTTGACCCCGCTAGATGAAGA- 3397
 QY 4246 GAAGCTCGCTCGATGATCTCCACCATGAGCAGCGGCGATGATGTC-----GA 4293
 Db 3398 AAGCCCTTGAACGATGCCAGGAGGTAGCACCAGGCGCTGGGACACATCACTCGCC 3457
 QY 4294 CCAGATGAGCGGACATTCACCGGCTTTTTCGAATTTTACCTTCGCTGCTTCTTCGACGA 4353
 Db 3458 CATAACAGCATCCTGAGCAGCGCATTTTACGATTCACCTCCGAGAGGTCTTTGATGA 3517
 QY 4354 CAA 4356
 Db 3518 CAA 3520
 RESULT 8
 AAA59129

Query Match 9.8%; Score 558.6; DB 20; Length 5276;
 Best Local Similarity 64.1%; Pred. No. 6.9e-108;
 Matches 964; Conservative 4; Mismatches 493; Indels 42; Gaps 7;

QY 2884 GGGCGGACGAGGATCCACCGGACACCGCTCTCTGTCGACGAGCTGCAAGAGAT 2943
 Db 2030 GGGCGGACGAGGATCCACCGGACACCGCTCTCTGTCGACGAGCTGCAAGAGAT 2089
 QY 2944 CAAGCAGCCCTCGCCCATCCCGGAGCGCAAGAGATCTCCAGGAGAGGAGGAGGAGCG 3003
 Db 2090 GATTCAGACCTCTCCCGTCCGCGCAAAACAGC-----AGAGACGAGAGAA 2140
 QY 3004 CAAGGAGTTCGGATCAGGAGTACTCGCCGCAAGCTTCAAGTTCTTTCATGAGCAGCA 3063
 Db 2141 GAGAGAGTCAAGCAGTCAAGAGCTACTCGCCATAGCGCTTTAAGTTCTTTCATGAGCAGCA 2200
 QY 3064 CATAGAGAGCTGATCAAGTCTGATCCGACGCGACGATCGCAAGAGTCAAGTGGAGAA 3123
 Db 2201 CGTGGAGAGTCTATCAAAACCTTACAGCAGAGAGTTAACCAGGAGCTGCGCTGGAGCA 2260
 QY 3124 GGAGATGACAAAGTGGGACTCGCCGATCAGACCCAAATCGAGATGAGGAAATGCTGAA 3183
 Db 2261 AGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATGCGGAGATCTCTTA 2320
 QY 3184 CCAAAGGAGAGCACTACATGATGATGAGCGGCGCAAGATGAGCAGAGATGTTCTGT 3243
 Db 2321 CCAGAAAGAGTCTAATTTACAAAGGTTTAAAGAGGGCGCAAGATGGACAAAGTCTATGTTGT 2380
 QY 3244 CAAATGAAGCCCATTTGGAGTGGTGTGATTTGGGAGGTAACTGCTGTGAGCAAAATCGA 3303
 Db 2381 CAGATCAAAACCTGGGAGTGGTGTGCTTTGGAGAGTGTGCTGCTTGTAGGTGGA 2440
 QY 3304 TACCTCGAACCATTTGTATCGGATGAAACCCCTCGGAAAGCGGACGTTCTCAAGCGGAA 3363
 Db 2441 ---CACTACGCGCTGTGACGCCATGAAGACCCCTAAGGAAAAAGGATGCTCTGAACCGGAA 2497
 QY 3364 TCAGGTGGCACAGCTGAGCGCCGAGGAGGATATCTTCGCGGAGCGGACCAATTAAGTGGGT 3423

2213	CGTGAGAAATGTCTCAAAAACCTTACCAGCAGAAGGTTTAAACGGAGGCTGCAGCTGGAGCA	2217
3124	GGAGATGCACAAAGTGGGACTGCCGATACAGACCCAAATCGAGATGAGAAAATGCTGAA	3183
2273	AGAAATGCCAAAGCTGACTCTGTGAAGCTGACGAGCAGATCGGAAGATCTCTTA	2332
3184	CCAAAGGAGAGCAACTACATTCTGATTGAAGCGGCCAAGATGACAAGAGCATGTTGGT	3243
2333	CCAGAAAGAGTCTAATTACAACAGTTAAAGAGGGCCAAAGATGACAAGTCTATGTTGT	2392
3244	CAAACTGAAGCCCATTTGAGTGGGTGCATTTTGGCGAGGTAACTGCTGCTGAGCAAAATCGA	3303
2393	CAAGATCAAAACCTTGGGGATCGGTGCTTTTGGAGAAGTGTGGCTTCTGTTAAGGTGGA	2452
3304	TACCTCGAAACCATTTGTATCGGATGA AAAACCTTCGGAAAAGCGGACGCTTCTCAAGCGAA	3363
2453	---CACTCAGCCCTGTACGCCATGAAGACCTTAAGAAAAAGGATGCTCTGAACCGAA	2509
3364	TCAGGTGCGACACGTGAAGCGCGAGAGGGATATCTCCGCGAAAGCGACAAATAACTGGGT	3423
2510	TCAGGTGCCACGCTCAAGCGCGAGAGGGACATCCTGCCCGAGCAGACAATCAGTGGGT	2569
3424	GGTGAAGTTGTACTACACTTCCAGGACAAAGGATAATCTGTACTTTGTGATGACACTACAT	3483
2570	GGTCAAACTCTACTACTCTTCCAAAGCAAAAGACAGCGCTGTACTTTGTGATGGACTACAT	2629
3484	ACCAGGTGTGATGTGATGTCGTGCTCATCAAACTGGGCATTTTCGAGGAGCAACTGGC	3543
2630	CCCTGGTGGGACATGATGACCTTGCTGATCCGGATGGAGGTCTTCCTGAGCACCTGGC	2689
3544	CAGATTCTACATCGCGGAGGTCACTGCGCGTGGACAGCGTTTCACAAAATGGGCTTCAT	3603
2690	CGGTTCTTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCAAAGATGGCTTCAT	2749
3604	TCACAGACATCAAGCCTGACACATCTCATCTGATAGGAGCGGACACATAAAGCTCAC	3663
2750	CCACGAGACATCAAGCCTGATAAATTTGTATAGATCTGGATGGTCACATTTAACTCAC	2809
3664	CGACTTTGGCTGTGACGGGATTTCCGATGAGCAGCACAACTCGAAGTACTACCAGAGAGA	3722
2810	AGATTTCCGCTCTGCACTGGGTTCAAGTGGACTCACANTTCCAAATATTACCAGAAAGG	2869
3723	-----ACGGCAATCACTCGCCAGGACTCGATGGAGCCCTGGGAGGAATACTCCGA---	3774
2870	GAGCCATGTACAGACGAGACATGGAGCCCGAGGACCTCTGGGATGATGTGCTAACTG	2929
3775	---GAAACGGACGAAGCCCACTGCTGGAGAGCGCAGGATGGCGGATCACCAAAGAGT	3831
2930	TGSGTGTGGGACAGGCTGAAGACCCCTAGAGAGAGGGCGGAAAGCAGCAGAGGTG	2989
3832	CCTGGCCCACTCGCTGGTGGGACCCCGAACTACATAGCTCCGAGGTGCTGCAGAGRAG	3891
2990	CCTGGCACATTCATGTTGGGGACTCCAACTACATCGCACCCGAGGTGCTCTCCGCAA	3049
3892	TGGGTACACGAGCTGTCCGACTACTCGAGCGCTGGGCGTCACTCTTAYGAGATGCTGGT	3951
3050	AGGGTACACTCAACTCTGTACTGGTGGAGTTGGAGTGATTCTCTTCGAGATGCTGGT	3109
3952	GGTCAAGCGCCCTTTCTGGGCCACAGCTCGCTGGAACCGACAAAGGTCTATCAACTG	4011
3110	GGGGCAGCGCCCTTTTGGCACCTACTCCCCACAGAACCCAGCTGAAGGTGATCAACTG	3169
4012	GGAGAAACCTGCATATTCGCCCGAGGCGGAGTTATCCCGCGAGGTACGACACTTGAT	4071
3170	GGAGAAACAGCTTCCATATTCAGCCCAAGGTGAAGCTGAGCCCTGAGGCCAGGACCTCAT	3229
4072	AAGGAGGCTCTGTGCTGGCTGACAAAGCGGCTGGGCAAGAGCGGTGGACAGG---TCAA	4128
3230	CACCAAGCTGTGCTCTCCGACAGACACCGCTTGGGCGGAATGGGCGCGATGACCTGAA	3289
4129	GAGCCACGACTTCTTCAAGGGCATCGACTTTGC---GGACATCGGAAGCAGAAACGGCC	4185
3290	GGCCCCACCCCTTCTTACGCGCAATTGACTTCTCCAGTGACATCCGAAGCAGCCACGCC	3349

QY 4186 CTACATACCGGAATCAACACCCACCCACGACACATCCAACTTTGATCCGTCGTCGGA 4245
 DB 3350 CTACGTTCCACCATCAGCCACCCACCCACGACACCTCGAATTTGACCCCGTAGATGAAGA 3409
 QY 4246 GAAGTCGCTCGATGACTCCACCATGATGACGCGCGGATGATGTC-----GA 4293
 DB 3410 AAGCCCTTGAACAGATGCGAGGAGGTAGCACCACCAAGGCTTGGGACACACTCACCTCGCC 3469
 QY 4294 CCAGATGACCCGACGCTTCCACGCGCTTTTTCGAATTTTACCTTCCGTCGCTTCTTCGACGA 4353
 DB 3470 CAATACAGATCTTGACGACGCAATTTACGAATTCACCTTCCGAGGTTCTTTGATGA 3529
 QY 4354 CAA 4356
 DB 3530 CAA 3532

RESULT 9
 AAA59130
 ID AAA59130 standard; DNA; 5486 BP.
 AC AAA59130;
 DT 07-NOV-2000 (first entry)
 DE DNA encoding a tumour suppressor protein hGHITS2.

Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 diabetic nephropathy; cardiopathy; tumour; breast cancer;
 renal adenocarcinoma; colorectal cancer; leukaemia; ss.

Homo sapiens.

Key Location/Qualifiers
 FH 387..3653
 CDS /*tag= a
 FT /product= "tumour suppressor protein"

EP1022333-A1.

26-JUL-2000.

07-OCT-1999; 99EP-0119199.

25-JAN-1999; 99JP-0016223.

(JCRP-) JCR PHARM CO LTD.

Koga J, Kono K, Zolotaryov FN;

WPI: 2000-516013/47.

P-PSDB; AAB07664.

New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 nephropathy or cardiopathy

Claim 1; Page 28-39; 59pp; English.

The present sequence encodes a human growth hormone inhibited tumour
 suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 antineoplastic pharmaceutical preparation. Probes for the hGHITS
 DNA sequences can be used in diagnostic pharmaceutical preparations.
 The diagnostic pharmaceutical preparations can be used for examining
 expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 tumours including breast cancer, renal adenocarcinoma, colorectal
 cancer, and leukaemia. Antibodies against the proteins can be used in
 a diagnostic pharmaceutical preparation for examining expression of
 a tumour suppressor gene.

SO Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;
 Query Match 9.8%; Score 558.6; DB 21; Length 5486;
 Best Local Similarity 64.1%; Pred. No. 7e-108;
 Matches 964; Conservative 4; Mismatches 493; Indels 42; Gaps 7;
 QY 2884 GGGCGGACGACGATCCACCGGACACCGCTCCCTCGTCGACACGATCCAGAAAGAT 2943
 DB 2042 GGGCGGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2101
 QY 2944 CAAGCAGCCCTCGCCATCCCGGAGCGCAAGAAGATCTCCAGGAGAGGAGGAGCG 3003
 DB 2102 GATTGAGACCTCTCCCGTTCCCGTCCGCGCAAAACAGC-----AGAGACGAGAGAA 2152
 QY 3004 CAAGGAGTTCGCGATCAGGAGTACTCGCCGCAAGCCTTCAAGTTCTTCATGGAGCAGCA 3063
 DB 2153 GAGAGAGTCAGCGCATCAAGAGCTACTCGCCATACGCTTTAAGTTCTTCATGGAGCAGCA 2212
 QY 3054 CATAGAGAACGTCATCAAGTCGTCGCGGACGACGATCGCAAGATCAAGTCGAGAA 3123
 DB 2213 CGTGAGAGTGTCAATCAAACTTACCAGCAGAAAGGTTAACCGGAGGCTGACGCTGGAGCA 2272
 QY 3124 GGAGATGCAAAAGTGGGACTGCGCGATCAGACCCAAATCGAGATGAGGAAATCTCTGAA 3183
 DB 2273 AGAAATGCCAAAGCTGGACTCTGTGAAGCTGAGCAGCAGATCGGAAGATCTCTTA 2332
 QY 3184 CCAAAGGAGAGCAACTACATTCGATTGAAGCGCGCCCAAGATGACAGAGCATGTTTCGT 3243
 DB 2333 CCAGAAAGAGTCTAATTACAAAGGTTAAAGAGGCGCAAGATGAGCAAGTCTATGTTGT 2392
 QY 3244 CAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGTAACTGCTGTGAGCAAAATCGA 3303
 DB 2393 CAAGATCAAAACCCCTGGGATCGTCTTGGAGAGTGTGCTTCTTGTAGGTGGA 2452
 QY 3304 TACCTCAACCATTTGATGGATGAAACCTCGGAAAGCGGACCTTCTCAAGCGGAA 3363
 DB 2453 ---CACTACGCGCTGTACGCCATGAAGACCTAAGGAAAGAGATGCTCTGAACCGGAA 2509
 QY 3364 TCAGTGGCACACGTCGAAGCGGAGAGGATATCTCGCGGAGGCGGACATCTCGGCGGAGGAGTCTTCCCTGAGCAGCCTGGC 2689
 DB 2510 TCAGTGGCGGACGTCGAAGCGGAGAGGAGTCTTCCGCGGAGGAGGAGTCTTCCCTGAGCAGCCTGGC 2689
 QY 3424 GTGAAGTTGTACTACAGCTTCAGGACAAAGGATTAATCTGTACTTTGTGAGTACTACAT 3483
 DB 2570 GGTCAAACTCTACTACTCTTCCAAAGACAAAGACAGCTCTGACTTTGTGAGTACTACAT 2629
 QY 3484 ACCAGTGGTGTATGTCGCTGCTCATCAAACTGGGATTTTCGAGAGGAGTGGC 3543
 DB 2630 CCCTGGTGGGACATGATGAGCCTGTGATCCGCGATGAGGTTCTTCCCTGAGCAGCCTGGC 2689
 QY 3544 CAGATTCTACATCGCGGAGTCACTTCGCGCGGAGGAGTTCACAAAATGGCTTCAT 3603
 DB 2690 CCGGTTCTACATCGCAGAGTGTACTTTGGCCATGAGAGTGTCCCAAGATGGCTTCAT 2749
 QY 3604 TCACAGAGATCAAGCCTGACAACTACATCATGATGAGGACGACACATAAAGCTCAC 3663
 DB 2750 CCACCGAGATCAAGCCTGATAAATTTGATAGATCTGATGGTTCACATTAACCTCAC 2809
 QY 3664 CGACTTTGGCTCTGACGAGGATTCGATGGACGACCACTCAAAAGTACTACAGAGAGA - 3722
 DB 2810 AGATTTCCGGCTCTGCACTGGGTTTCAGGTGGACTCAATATTCAAAATATTCACAAAGG 2869
 QY 3723 -----ACGCAATCACTTCGCGCGGAGGAGTCTGATGGAGCCTGGGAGGATATCTCCGA --- 3774
 DB 2870 GAGCCATGTACAGACGACAGCATGAGCCCGACGACCTCTGGATGATGTCTTAATC 2929
 QY 3775 ---GAGACGACGAGCCACCTGCTGTGAGAGGCGGAGTGGCGGATCAACAAAGT 3831
 DB 2930 TCGGTGTGGGAGCAGGCTGAAGACCTTAGAGCAGAGGCGGCGGAGCAGCAGAGGTG 2989
 QY 3832 CTTGGCCCACTCTGCTGGTGGGACCCGCACTACATAGTCCCGAGGCTCTCGAGAGAG 3891
 DB 2990 CTTGGCAGATCTACTGTTGGGAGTCTCAAACTACATCGCACCGGAGGTGCTCTCTCCGCA 3049

QY 3892 TGGTACACGACGAGTGTGCGACTACTGAGCGTGGCGTTCATCTCTAYGAGATGCTGGT 3951
 DB 3050 AGGTACACTCACTGTGACTGTTGGAGTGTGGAGTATCTCTTCAGATGCTGGT 3109
 QY 3952 GGGTACGCGCCCTTTCTGCGCAACAGTCCGCTGGAAACCAACAAAGTTCATCAACTG 4011
 DB 3110 GGGGCGAGCCGCCCTTTTGGCACTACTCCACAGAAACCCAGCTGAAGTGTATCAACTG 3169
 QY 4012 GGAGAAACCTCATATTCGCGCGAGCGCGGAGTTATCCGCGAGGTACGACTGTAT 4071
 DB 3170 GGAGAACACCTCCACATTCACGCCAGGTGAAGCTGAGCCCTGAGGCCAGGACCTCAT 3229
 QY 4072 AAGGAGCTCTGTCGCTGCGCTGACAGCGCTGGCAAGAGCGTGGACGAGG---TCAA 4128
 DB 3230 CACCAAGCTGTGCTGCCAGCACACCGCTGGGCGGAATGGGCGCATCACCTGAA 3289
 QY 4129 GAGCCAGACTTCTCAAGGGCATGCACTTTGCG---GGACATCGGAAAGCAAGCGCC 4185
 DB 3290 GGGCCACCCCTTCTGAGCGCCATTTGACTTCTCCAGTGACATCCGGAAGCAGCCGCC 3349
 QY 4186 CTACATACCGGAATCAAGCACCCACGACACATCAACTTTGATCCCGTGTATCCGGA 4245
 DB 3350 CTAGCTTCCACCATCAGCCACCCCGAGGACCTCGAATTCGACCCCGTAGATGAAGA 3409
 QY 4246 GAAGCTGCGCTCGAATCACTCCACCATGAGCAGCGCGATGATGTC-----GA 4293
 DB 3410 AAGCCCTTGAAGCATGCCAGCAAGGTAGCACCAAGCGCTGGACACACTCACCTCGCC 3469
 QY 4294 CCAGATGACCGCAYTTCACGGCTTTTCGAATTTACCTTCGCTGCTTTTCGACGA 4353
 DB 3470 CAATAACAGCATCTCTGAGCAGCATTTTACGAATTCACCTCCGGAAGTTCCTTGATGA 3529
 QY 4354 CAA 4356
 DB 3530 CAA 3532

RESULT 10
 ID AAZ61160
 XX AAZ61160 standard; cdna; 1961 BP.
 AC AAZ61160;
 XX 30-MAY-2000 (first entry)
 XX cdna SSI771A encoding a partial protein kinase.
 DE Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH 3.1769
 FT CDS /*tag= a
 FT /note= "partial sequence"
 XX
 XX WO200008180-A2.
 PN
 XX 17-FEB-2000.
 XX
 XX 03-AUG-1999; 99WO-US17630.
 XX
 XX 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0099972.
 XX
 XX (IMMUNEX CORP.
 XX Virca GD, Bird TA, Anderson DM, Marken JS;
 PI WPI; 2000-195584/17.
 DR P-PSDB; AAY69162.

XX New human kinase polypeptides and polynucleotides used as molecular
 PT weight markers and as controls for peptide fragmentation
 XX
 PS Claim 1; Page 9; 60pp; English.
 XX
 CC The present sequence encodes a partial polypeptide which has kinase
 CC activity. The kinase polynucleotides can be used to express the
 CC polypeptides, and as probes to identify nucleic acids encoding
 CC proteins having kinase activity. The kinase polypeptides and
 CC fragmented polypeptides are used as molecular weight and isoelectric
 CC focusing markers, and as controls for peptide fragmentation. They also
 CC have a number of therapeutic uses as kinases play a central role in
 CC cellular signal transduction. The polypeptides could also be used to
 CC identify binding partner proteins. The polypeptides can be used as
 CC a reagent to identify any proteins that the polypeptide regulates, and
 CC proteins with which it might interact. The polypeptides may also be
 CC used for preparation of antibodies. The antibodies can be used in
 CC assays to detect the presence of the protein, and to purify the protein
 CC by immunoaffinity chromatography.
 XX
 SQ Sequence 1961 BP; 504 A; 550 C; 539 G; 368 T; 0 other;
 Query Match 9.7%; Score 557; DB 21; Length 1961;
 Best Local Similarity 64.1%; Pred. No. 9.9e-108;
 Matches 963; Conservative 4; Mismatches 494; Indels 42; Gaps 7;
 QY 2884 GGGCGGACGACGAGTATCCACCGCACCGCTCTCTCGTCACAGCTGCAGAGAT 2943
 DB 158 GGGCGGACGACGAGTATCCACCGCACCGCTCTCTCGTCACAGCTGCAGAGAT 2943
 QY 2944 CAAGCAGCGCTCCGCCATCCGCGAGCGCAAGAGATCTCAAGGAGAGGAGGAGCG 3003
 DB 218 GATTGAGACTCTCCCGTTCGCCGCAAAACAGC-----AGAGACGAGAGAA 268
 QY 3004 CAAGGAGTTCCGATCAGGAGTACTCGCGCAAGCTTCAAGTCTTCTTATGAGAGCA 3063
 DB 269 GAGAGAGTCAAGCATCAAGAGTACTCGGCATACGCTTTAAGTCTTCTTATGAGAGCA 328
 QY 3064 CATAGAGACGATCAAGTCTGATCGCGCACGCTATCGCAAGATCAAGTCTGAGAA 3123
 DB 329 CGTGGAGATGTCTCAAAACCTTACAGAGAGAGTAAACCGGAGCTCGAGTGGAGCA 388
 QY 3124 GGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGAAATGCTGAA 3183
 DB 389 AGAATGGCCAAAGCTGGACTCTGTGAAGTGGAGAGGAGAGTGGAGAGTCTCTTA 448
 QY 3184 CCAAGAGGAGACAACTACATTCGATTGAAGCGCCCAAGATGGACAAGAGATGTTGCT 3243
 DB 449 CCAGAAAGAGTCTAATTACAAAGTTAAAGAGGGCCAAAGTGGACAAGTCTATGTTGT 508
 QY 3244 CAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAAACGCTGAGCAAAATCGA 3303
 DB 509 CAAGATCAAAACCCCTGGGATCGTGCCTTTGGAGAGAGTGTGCTTGTAAAGTGA 568
 QY 3304 TACCTCGAACCACTTTCTATGCGATGAAACCCCTCGGGAAGCGAGCTCTCAAGCGAA 3363
 DB 569 ---CACTCAGCCCTGTAGCCATGAAGACCCCTTAAGGAAAGAGATGTCTGAACCGAA 625
 QY 3364 TCAGTGGGACACGCTGAAGGCCGAGAGGATATCTCCGGAAGCGGACAACTAGTGGT 3423
 DB 626 TCAGTGGGACACGCTGAAGGCCGAGAGGACATCTCTGGCGGAGGACAACTAGTGGT 685
 QY 3424 GGTGAAGTGTACTACAGCTTCCAGGACAGAGTATCTGTACTTGTGATGACTACAT 3483
 DB 686 GGTCAAACTCTACTACTCTCTCCAGCAAGAGACAGCTGTACTTGTGATGACTACAT 745
 QY 3484 ACCAGTGTGATCTGATGCTGCTCATCAAACTGGGCTTTTCGAGAGGAACTGCG 3543
 DB 746 CCTGTGGGGACATGATGAGCTGTCTGATCGGATGGAGGTCTTCCCTGAGACCTGCG 805
 QY 3544 CAGATTCTACATCGCGGAGGTACCTGCGCCGCTGGAGAGGTTTCAAAATGGCTTCAT 3603

806	CGGGTCTTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCACAAGATGGGCTTCAT	865
3604	TCACAGAGACATCAAGCCTGCAGAACTACTCATCGATAGGAGCGGACACATAAAGCTCAC	3663
866	CCACCGAGACATCAGCCTGATACATTTTGCATAGATCTGGATGGTCACATTTAAACTCAC	925
3664	CGACTTTGGCTGTGCACGGGATTCGATGGACGCACAATCGAAGTACTTACACAGAGA-	3722
926	AGATTTGGGCTCTGCATCGGTGTGAGTGACTCACAATTCCAAATATTACCAAGAGG	985
3723	-----ACGCAATCACTCGCGCAGGACTCGATGAGCGCCCTGGAGAGAACTACTCGA---	3774
986	GAGCCATGTCACAGAGGACAGCATGGAGCCCGGACCTCTGGGATGATGTGCTTAACGTG	1045
3775	---GAAGCGACCGAAGCCACCGCTGCTGGAGAGCGACGGATCGCGGATCACCAAGAGCT	3831
1046	TCGGTGTGGGACAGGCTGAAGACCTTAGACAGAGGCGCGGAAGCAGCACAGAGGTG	1105
3832	CCTGGCCCACTCGCTGTGGGCACCCCGAACTACATAGCTTCCGAGGTCGTCGAGAGRAG	3891
1106	CCTGGCACAATTCAGTGTGGGGACTCCAAACTACATCGACCCGAGGTGCTCTCGCGAA	1165
3892	TGGGTACAGCAGCTGTGCGCACTACTGGAGCGTGGGCGTCATCTTAYGAGATGCTGCT	3951
1166	AGGGTACACTCAACTCTGTGACTGGTGGAGTGTGGAGTGATTTCTCTCGAGATGCTGT	1225
3952	GGGTACAGCGCCCTTTCTGGGCCAACAGTTCGCTGGAAACGCAACAAAGGTCACTAAGT	4011
1226	GGGGAGCGCGCCCTTTTGGCACCTTACTCCACAGAAACCCAGCTGAAGGTGATCACTG	1285
4012	GGAGAAACSCGTGCATATTCGCGCGCAGGCGGAGTTATCCGCGAGGCTACGACTTGAT	4071
1286	GGAGAACAGCTCCACATTCACGCCCAAGTGAAGCTGAGCCCTGAGGCCAGGACCTCAT	1345
4072	AAGGAGGCTCTGTGCTCGGCTGCAGACGGCTGGGCAAGAGGCTGGAGCAGG---TCAA	4128
1346	CACCAAGCTGTCTGCTCCGCGACAGCACCGCTGGCGGGAATGGGCGGAGTGAACCTGA	1405
4129	GAGCCACAGCTTCTTCAAGGGATGAGCTTTGC---GGACATCGGAGCAGAAAGCGCC	4185
1406	GGCCCAACCCCTTCTTCAAGCGCATTTGACTTCTCAGTGGATCCGGAAGCATCCAGCCC	1465
4186	CTACATACCGGAATCAAGCACCCACGAGACACATCCAACTTTGATCCCGTGGATCCGGA	4245
1466	CTACGTTCCCAACCATCAGCCACCCCATGGACACCTCGAATTCGACCCGTAGATGAAG	1525
4246	GAACTGCGCTCGAATGACTCCACCATGAGCAGCGCGGATGATGTC-----GA	4293
1526	AAGCCCTTGGAAACGATGCGCAGGAGTAGCACCAAGGCGCTGGGACACACTACCTCGCC	1585
4294	CCAGATGACCCGACACTTCCACGGCTTTTTCGAATTTTACCTTCGCTGCTTCTTCGACGA	4353
1586	CATAACAAGCATCTGAGCAGCAATTTTACGAATTCACCTTCCGAGGTTCTTTGATGA	1645
4354	CAA 4356	
1646	CAA 1648	

RESULT 11
AA11765
ID AA11765 standard; DNA; 3533 BP.

RESULT 11

AAI71765

ID AAI71765 standard; DNA; 3533 BP.

XX
XX
/TTH;
DTAAI73
AC

XX XX

DT 18-J2

XX

DE Human

XX

KW Human

XX

Homo
SO
vv

XX

.

JP2001231565-A.
 28-AUG-2001.
 18-FEB-2000; 2000JP-0041818.
 18-FEB-2000; 2000JP-0041818.
 (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.
 (NOJIT/) NOJIMA H.
 WPI: 2001-650995/75.
 P-PSDB; AAG78992.
 Novel human Lats2 gene useful for diagnosing cancer -
 Claim 1; Page 17-18; 34pp; Japanese.
 The present sequence is the coding sequence for human cancer-inhibiting
 gene Lats2. Lats2 is a kinase and can be used for gene diagnosis of
 cancers.
 Sequence 3533 BP; 831 A; 1096 C; 1011 G; 595 T; 0 other:
 Query Match 9.7%; Score 557; DB 22; Length 3533;
 Best Local Similarity 64.1%; Pred. No. 1.3e-107;
 Matches 963; Conservative 4; Mismatches 494; Indels 42; Gaps 7;
 QY 2884 GGGCGGCGAGCAACGGATCCACCGGCACACCGCCCTCTCTGTCGACAGCTGCACAGAGAT 2943
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1653 GGGCGGCGAGCAACGGATCCACCGGCACACCGCCCTCTCTGTCGACAGCTGCACAGAGAT 1712
 QY 2944 CAAGCAGCCTCGCCCATCCGGAGCGCAAGAACATCTCCAAGGAGAGGAGGAGGAGCG 3003
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1713 GATTCAGACCTCTCCCGTTCCTGTCGCGCAAAACAGC-----AGACGAGAGGAA 1763
 QY 3004 CAAGGATTCGCGATCAGGCGAGTACTCCGCGCAAGCCTTCAAGTCTTCTATGGAGCAGCA 3063
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1764 GAGAGAGTCAGCATCAAGAGCTACTCGCCATACGCCCTTTAAGTCTTCTATGGAGCAGCA 1823
 QY 3064 CATAGAAGCTGATCAAGTCGTATCCCGCAGCGCACTATCGCAAGAGATCAGCTGGAGAA 3123
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1824 CGTGGGAATGTCTATCAAACTTACCAGCAAGAGTTAACCGGAGGCTGCAGCTGGAGCA 1883
 QY 3124 GGAGATCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATGCTGAA 3183
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1884 AGAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATCGGGAAGATCTCTA 1943
 QY 3184 CCAAGAGAGCAACTACATTCGATTCAAGCGCGCCCAAGATGACAGAGACATGTTCTCGT 3243
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 1944 CCAGAAAGAGTCTAATTACACAGGTTAAAGAGGGCCCAAGATGACAGATCTATGTTGT 2003
 QY 3244 CAAACTGAAGCCCATTTGGAGTGGGTGCTATTGGCGAGGTAACTGGTGAAGCAAAATCGA 3303
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 2004 CAAGATCAAAACCTGGGATCGTGCTTTGGAGAGTGTGCTTCTTGTGAAGTGA 2063
 QY 3304 TACCTGAACCAATTTGATGGATGAACACCTCGGGAAGCGGACGTTCTCAAGCGGAA 3363
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 2064 ---CACTCAGCCCTGTACGCCATGAAGACCTTAAGGAAAAGGATGTCTTGAACCCGAA 2120
 QY 3364 TCAGGTGCGACACTGAAGCCGAGAGGATATCCTCGCGAAGCCGACAAATAACTGGGT 3423
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 2121 TCAGGTGGCCACCTGAAGCCGAGAGGACATCTTGCCGAGCGACACAATGATGGGT 2180
 QY 3424 GGTGAAGTTGTACTACAGCTTCCAGGACAGGATAATCTGTACTTTGTGATGGACTACAT 3483
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 2181 GGTCAAACTCTACTACTCTTCCAAGACAAAGACAGCCTGTACTTTGTGTGAGACTACAT 2240
 QY 3484 ACCAGTGGTGATCTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGGACTGCG 3543
 Db ||||| - - - - - ||||| - - - - - ||||| - - - - - ||||| - - - - - |||||
 2241 CCCGTGGGACATGATGAGCCTGTCTATCCCGATGGAGTCTTCCCTGAGCACCCTGGC 2300
 QY 3544 CAGATTCTACATCCCGAGGTACCTCGCCGCTGGACAGCGCTTCACAAAATGGGCTCAT 3603

Db 2301 CCGGTTCTACATCGCAGAGCTGACTTGGCCATTGAGAGTGTCACAAAGATGGGTTTCAT 2360
 Qy 3604 TCACAGAGACATCAAGCCTGACAACTACTCATCATAGGAGCGGACACATAAAGCTCAC 3663
 Db 2361 CCACCGAGACATCAAGCCTGACAACTTTGATGATCTGGATGTTGATCTAAACTCAC 2420
 Qy 3664 CGACTTTGGCTGTGTCACGGGATTCGATGGAGCGACAACTCGAACTACTACAGAGAGA- 3722
 Db 2421 AGATTTCGGCTCTGCAGTGGTTCAGGTGGACTCACATTCARATATACAGAAAGG 2480
 Qy 3723 -----ACGGCAATCACTCGCCAGGACATCGATGAGCCCTGGGAGGAATCTCCGA --- 3774
 Db 2481 GAGCCATGTCACAGAGACAGCATGAGCCAGCCAGCCCTCTGGATGATGTTCTAACTG 2540
 Qy 3775 ---GAACGACCAAGCCACCGTCTGGAGAGCGCAGGATGCGGATCAACCAAGAGT 3831
 Db 2541 TCGGTGTGGGAGAGCTGAAGACCTAGAGCAGAGGGCGGAGACACACAGAGGTC 2600
 Qy 3832 CCGTGCACACTCGCTGGGACCCGAACTACATAGCTCCGAGGTGCTGGAGAGAG 3891
 Db 2601 GCTGCACATTCAGTGGGAGCTCCAACTACATCGCACCGAGGTGCTCTCGGCAA 2660
 Qy 3892 TGGGTACACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3951
 Db 2661 AGGGTACACTCACTCTGTGACTGTGGAGTCTGGAGTGATCTCTCGAGATGCTGT 2720
 Qy 3952 GGGTACGCGCCCTTCTGGCCACAGTCGCTGGTGGAAAGCAACAAAGGTTCATCACTG 4011
 Db 2721 GGGGAGCGCGCCCTTTTGGCACCTACTCCACAGAAACCCAGCTGAAGGTGATCAACTG 2780
 Qy 4012 GGAGAAACSCGTGATATTCGCGCGCAGCCGAGTTATCCCGAGGCTACGACTTGAT 4071
 Db 2781 GGAGAACGCTCCACATTCAGCCAGGTGAAGCTGAGCCCTGAGCCAGGAGCTCAT 2840
 Qy 4072 AAGGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4128
 Db 2841 CACCAAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2900
 Qy 4129 GAGCCAGCACTTCTCAAGGGATGCACTTTGC---GGACATCGGGAACCAAGCGCC 4185
 Db 2901 GGGCCACCCCTTCTCAGCGCATGACTTCTCCAGTGACATCCCGGAGCAGCCAGCCCC 2960
 Qy 4186 CTACATACCGGAAATCAAGCACCACGACACATCCAACTTTGATCCCGTGATCCGA 4245
 Db 2961 CTACGTTCCACCATCAGCCACCCATGACACCTCGAATTCGACCCGCTAGATGAAGA 3020
 Qy 4246 GAAGCTGCGCTCGAATGCTCCACCATGAGCAGCGCGGATGATGTC-----GA 4293
 Db 3021 AAGCCCTTGGACGATGCCAGGGAAGGTAGCACCAAGGCTGGACACACACTCACCTCGCC 3080
 Qy 4294 CCAGATGACCGCACYTTCCAGCGCTTTTTCGAATTTACCTTCGCTCGCTTCTTCGACGA 4353
 Db 3081 CAATAACAGCATCTGAGCAGCATTTTACGAATTCACCTTCGGAAGGTTCTTTGATGA 3140
 Qy 4354 CAA 4356
 Db 3141 CAA 3143

RESULT 12

AAZ61158
 ID AAZ61158 standard; cDNA; 1498 BP.
 XX
 AC AAZ61158;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE cDNA S5171 encoding a human protein kinase.
 XX
 KW Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX

OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 2..1498
 FT /*tag= a
 FT /*product= "kinase"
 XX
 PN WO200008180-A2.
 PD 17-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17630.
 XX
 PR 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0099972.
 XX
 PA (IMMV) IMMUNEX CORP.
 XX
 PI Virca GD, Bird TA, Anderson DM, Marken JS;
 DR WPI: 2000-195584/17.
 DR P-PSDB; AAY69160.
 XX
 CC New human kinase polypeptides and polynucleotides used as molecular weight markers and as controls for peptide fragmentation
 PS Claim 1; Page 8-9; 60pp; English.
 XX
 CC The present sequence encodes a polypeptide which has kinase activity. The kinase polynucleotides can be used to express the polypeptides, and as probes to identify nucleic acids encoding proteins having kinase activity. The kinase polypeptides and fragmented polypeptides are used as molecular weight and isoelectric focusing markers, and as controls for peptide fragmentation. They also have a number of therapeutic uses as kinases play a central role in cellular signal transduction. The polypeptides could also be used as identifying binding partner proteins. The polypeptides can also be used as a reagent to identify any proteins that the polypeptide regulates, and proteins with which it might interact. The polypeptides may also be used for preparation of antibodies. The antibodies can be used in assays to detect the presence of the protein, and to purify the protein by immunoadfinity chromatography.
 XX
 SQ Sequence 1498 BP; 384 A; 417 C; 427 G; 270 T; 0 other;

Query Match. 9.4%; Score 536; DB 21; Length 1498;

Best Local Similarity 65.5%; Pred. No. 2.4e-103;

Matches 884; Conservative 3; Mismatches 433; Indels 30; Gaps 6;

Qy 2884 GGGCGGCGACGAGATCCACCGCACCGCTCTCTGTCGACCGCTGCAAGAAGAT 2943
 Db 160 GGGCGGCGACGAGATCCACCGCACCGCTCTCTGTCGACCGCTGCAAGAAGAT 2943
 Qy 2944 CAAGCAGCGCTCGCCCATCCCGAGCGCAAGAGATCTCCAAAGGAGAGAGGAGCG 3003
 Db 220 GATTCAGACCTCTCCCGTTCGCGCAAAACAGC-----AGAGACGAAGAGAA 270
 Qy 3004 CAAGGATTCGCGATCAGGAGTACTCGCCGCAAGCTTCAAGTCTTTCATGAGCAGCA 3063
 Db 271 GAGAGATCAGCATCAAGAGTACTCGCCATACGCTTTAAGTCTTTCATGAGCAGCA 330
 Qy 3064 CATAGAGAAGTGTCAAGTCTGATCGCCGCGCAGCTGATCGCAAGTCAAGTCTGAGAA 3123
 Db 331 CGTGGAGATGTATCAAAACCTACAGCAGAGGTTAAACCGGAGCTCGACTGGAGCA 390
 Qy 3124 GGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATTCGAGATGAGGAAATGCTGA 3183
 Db 391 AGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATCGCGAAGATCTCTA 450
 Qy 3184 CCAAAAGGAGAGCACTACTATTCGATTGAAGCGCCGCAAGATGGACAAAGCATGTTCTGT 3243
 Db 451 CCAGAAAGAGTCTAATTAACAACAGGTTAAAGAGGGGCAAGATGGACAAAGTCTATGTTGT 510

Qy 3244 CAAACTGAAGCCCATTTGGAGTGGGTGATTTGGCGAGTAAACGCTGTGTGACGAAAAATCGA 3303
 Db 511 CAAGATCAAAACCCCTGGGATCGTCTTTGGAGAACTGTGCCTTGTAAAGTGA 570
 Qy 3304 TACCTCAACCACTTTGATGGGATGAACCCCTCGGAAAGCGACGCTTCTCAAGCGGAA 3363
 Db 571 ---CACTCAGCCCTGTACGCCATGAGACCCCTAAGGAAAGGATGCTCTGAACCGGAA 627
 Qy 3364 TCAGTGGCACACCTGAAGCCGAGAGGGATATCTCGCGGAAGCCGACACATAACTGGGT 3423
 Db 528 TCAGTGGCCACACCTCAAGCCGAGAGGGATCTCTGGCGAGCAGACATGATGAGTGGT 687
 Qy 3424 GGTGAAGTTGATACAGCTTCCAGGACAAGGATAATCTGTACTTTGTATGGACTACAT 3483
 Db 688 GGTCAAACTCTACTACTCTCTCCAGACAAGACAGCCTGTACTTTGTATGGACTACAT 747
 Qy 3484 ACCAGGTGGTATCTGATGCTGCTCATCAAACTGGGATTTTCGAGGAGGAACCTGC 3543
 Db 748 CCCTGGTGGGACATGATGAGCCTGCTGATCCGGATGAGGTCTTCCCTGAGCACCTGGC 807
 Qy 3544 CAGATTCTATATCCCGAGGTACCTGGCGTGAGCAGGGTTCCACAAAATGGGCTTCAT 3603
 Db 808 CCGGTTCTATATCCCGAGGTACTTTGGCCATTGAGAGTGTCCACAAGATGGGCTTCAT 867
 Qy 3604 TCACAGAGACATCAAGCCTGACACATCTACTATCATGATAGGAGCGGACACATAAGCTCAC 3663
 Db 868 CCACCGAGACATCAAGCCTGATCAATCTTTGATAGATCTGGATGGTGCACATTAACCTCAC 927
 Qy 3664 CGACTTTGGCTGTGCACGGGATTCGATGACGACCAACTCGAAGTACTACACAGGA --- 3720
 Db 928 AGATTTCGGGCTGTGCATCTGGGTTTCAGGTGGACTCACAATTTCCAAATATTACCAAGAAGG 987
 Qy 3721 ---GAACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAGGA -----ATACTC 3771
 Db 988 GAGCATGTACAGACAGACACATGAGCCGACCTCTGGGATGATGTGTCTAATG 1047
 Qy 3772 CGAAGAGGACCGAAGCCACCGTGTGGAGAGCGGAGTGCAGATCCACCAAGAGT 3831
 Db 1048 TCGGTGTGGGACAGGCTGAAGACCTTAGAGAGGCGCGGAAGCAGCAGAGAGTG 1107
 Qy 3832 CCTGGCCACTCGCTGTGGGACCGGACCTACATAGTCCCGAGTGTGGAGAGAG 3891
 Db 1108 CTTGGACATCACTGTGGGAGCTCCAACTATATCGACCCGAGGTCTCTCCCGAA 1167
 Qy 3892 TGGGTACAGCAGCTGTGCGACTACTGGAGCGTGGCGCTCATCTTAYGAGATGCTGT 3951
 Db 1168 AGGTACACTCAACTCTGTGACTGTGGAGTGTGGAGTGTCTCTCGAGATGCTGT 1227
 Qy 3952 GGTTCAGCGCCCTTTCTGGCCACAGTCCGCTGGAACCGCAACAAAGTTCATCACTG 4011
 Db 1228 GGGCAGCGCCCTTTTGGCACCTACTCCACAGAACCCAGCTGAAGTGTATCACTG 1287
 Qy 4012 GGAGAAACCTGATATTCGCGCAGCGGAGTATTCGCGGAGGCTACGGACTGTAT 4071
 Db 1288 GGAGAACAGCTTCCACATTCAGCCAGGTGAAGCTGAGCCCTGAGGCGAGGACCTCAT 1347
 Qy 4072 AAGGAGCTGTGCTGCTCCGAGACCCAGCGCTGGGCAAGAGCGTGGACGAGG---TCAA 4128
 Db 1348 CACCAAGCTGTGCTCTCCGAGACCCAGCGCTGGGCGGAATGGGCGGATGACCTGAA 1407
 Qy 4129 GAGCCAGACTTCTTCAAGGCGATCGACTTTGC---GGACATGGGGAAGAGAGCGCC 4185
 Db 1408 GGGCCACCCCTTCTTCAAGGCGATGACTTCTTCAGTGTGACATCCGGAAGCATCCAGCCCC 1467
 Qy 4186 CTACATCCGGAATCAAGCACCCACCGA 4215
 Db 1468 CTACGTTCCCACTACGCCACCCCATGGA 1497

RESULT 13
 ABK43457
 ID ABK43457 standard; cDNA; 1912 BP.
 XX

AC ABK43457;
 XX 05-JUN-2002 (first entry)
 DT
 XX DNA encoding novel central nervous system protein #37.
 DE
 XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy; gene; ss.
 XX Homo sapiens.
 OS
 XX WO20015318-A2.
 PN
 XX 02-AUG-2001.
 PD
 XX 17-JAN-2001; 2001WO-US01332.
 PF
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216847.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224318.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.

181 AAGATGGACAAAGTCTATGTTTGTCAAGATCAAAACCCCTGGGATCGGTGCCTTTGGAGAA 240

QY 3281 GTACGCTGTGACCAAAATCGATACCTCGAACCAATTTGTATGGGTGAAACCCCTCGG 3340
DB 241 GTGTGCTTGTGTTGAAGTGA---CACTACAGCCCTGTACGCCATCAAGACCTAAG 297
QY 3341 AAAGCGGACGTTCTCAAGCGGAATCAGTGGCAGCACAGTGAAGCCGAGGATATCCTC 3400
DB 298 AAAAAGATGTCTTGAACCGGAATCAGTGGCCACCTCAAGGCCGAGAGGATCCTG 357
QY 3401 GCGAAGCCGACATAACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATAAT 3460
DB 358 GCGGAGCAGACATAGTGGTGGTCAAACTCTACTCTCCATCCAGACAGACAGC 417
QY 3461 CTGTACTTTGTATGAGTACATACAGAGTGGTGTATCTGATGCTGCTCATCAAACTG 3520
DB 418 CTGTACTTTGTATGAGTACATACCTTGGTGGGACATGATGAGCTGCTGATCCGATG 477
QY 3521 GGCATTTTCGAGGAGGAGTGGCAGATTTCTACATCCCGAGGTCACTGCGCGTGGAC 3580
DB 478 GAGTCTTCCCTGAGCAGCTGGCCCGGTTCTACATCCGAGAGTGACTTTGGCCATTTGAG 537
QY 3581 AGGCTTCCACAAATGGCTTCATTCACAGAGACATCAAGCTGACAACTACTCATCGAT 3640
DB 538 AGTGTCCACAGATGGCTTCATCCACCGAGACATCAAGCTGATACATTTTCATAGAT 597
QY 3641 AGGACGCGACATAAAGTCTACCGACTTTGGCTGTGACAGGATTCGATGGACGAC 3700
DB 598 CTGATGCTCACATTAACCTCACAGATTTGGGCTCTGCACCTGGTTCAGTGGACTCAC 657
QY 3701 AACTCGAAGTACTACGAGGAGAGCGCAATCACTCGC-----GCCAGGACTCGATGGAG 3754
DB 658 AATTCCAATATTACCAAGAGGAGCCATGTGACAGAGACAGCATGGAGCCCGACGAC 717
QY 3755 CCCTGGGAGGAAT-----ACTCGGAGAACCGGACCGCCACCGTGTGGAGAGCGCA 3808
DB 718 CTCTGGATGATGTCTACTGTCTGTGGTGGGACAGGCTGAGACCTTAGACGAGG 777
QY 3809 CGGATGCGGATCACCAAGAGTCTTGGCCACTCGCTGTGGGACCCCGCAACTACATA 3868
DB 778 GCGGGAAGCAGCAGAGAGTGTGGACATTCATCTGTGGTGGGATCCAACTACATC 837
QY 3869 GTCGCGGAGTGTGGAGAGTGGTACACGAGCTGTGGCAGTGTGGAGGCTGGGC 3928
DB 838 GCACCCGAGTGTCTTCCGCAAGGATACATCACTGTGAGTGTGGAGTGTGA 897
QY 3929 GTATCTCTTAYGAGATGTGTGGTGTGAGCGCCCTTTCTGGCAACAGTCCGCTGGAA 3988
DB 898 GTGATCTCTTCGAGATGTGTGGTGGGACCGCCCTTTTGGCACCTACTCCACAGAA 957
QY 3989 ACCGACAAAGTGTATCAACTGGGAGAAACSTGTGATATTCGCGGAGGCGGAGTTA 4048
DB 958 ACCAGTGAAGTGTATCAACTGGGAGAACACGCTCCACATTCAGCCAGGTGAAGCTG 1017
QY 4049 TCCGCGGAGGCTACGAGTGTATGAAGAGGCTGTGCGTGGCTGACAGCGGCTGGC 4108
DB 1018 AGCCCTGAGGCGGAGGACCTTCATCACCAGCTGTGTCTCCGACAGACCCCGCTGGG 1077
QY 4109 AAGAGGCTGGACGAGG---TCAAGAGCCAGGCTTCTTCAAGGCGCATTCGACTTTCG---G 4162
DB 1078 CGGAATGGGCGGATGACCTGAGAGCCCGCCCTTCTCAGCGCATTCGACTTCTCCAGT 1137
QY 4163 GACATCGGAGAGAGAGCGGCTTACATACCGGAATCAAGCACCCCGAGGACATCC 4222
DB 1138 GACATCGGAGAGGAGCGGCGGCTACGTTCCACCATCAGCCAGCCCGATGGACGCTCG 1197
QY 4223 AACTTTGATCCCTGTGATCGGAGAGGCTGCTGCAATGACTCCACCATGACGCGGC 4282
DB 1198 AATTTGACCCCTGTATGAAGAGCCCTTGGAGACGATGCCAGGAGGTAGCACCAG 1257
QY 4283 GATGATGTC-----GACCAGATGACCGGACATCTTCACCGCTTTTTCGAAATTT 4330
DB 1258 GCCTGGGACACACTCACTCGGCCCAATAACAAGCATCCTGAGCAGCATTTTACGATTC 1317
QY 4331 ACCTTCGCTGCTTCTTCGACGACAA 4356

DB 1318 ACCTTCGAAGGTTCTTTTGATGACAA 1343

RESULT 14

AAT42119

ID AAT42119 standard; cDNA; 3213 BP.

XX AAT42119;

XX 31-JAN-1997 (first entry)

XX M-lats gene encoding large tumour suppressor.

XX Mouse; m-lats gene; large tumour suppressor; fetal brain;

XX protein-serine/threonine-kinase; cell proliferation; antisense;

XX dominant-negative; cancer; degenerative disorder; trauma;

XX growth deficiency; therapy; antitumour; vulnary; diagnostic;

XX transgenic plant; transgenic animal; growth; senescence; ds.

XX Mus musculus.

XX OS

XX PH

XX FT

XX CDS

XX Key

XX Location/Qualifiers

XX I..2889

XX /*tag= a

XX /product= m-lats protein

XX WO9630402-A1.

XX XX

XX PD

XX 03-OCT-1996.

XX XX

XX PF

XX 26-MAR-1996; 96WO-US04101.

XX XX

XX PR

XX 27-MAR-1995; 95US-0411111.

XX XX

XX PA

XX (UYA) UNIV YALE.

XX XX

XX PI

XX Tao W, Wang W, Xu T, Yu W, Zhang S;

XX XX

XX WPI; 1996-455275/45.

XX DR

XX P-PSDB; AAW05179.

XX XX

XX PT

XX New isolated large tumour suppressor gene - used to develop prods.

XX for inhibiting cell proliferation or for enhancing proliferation

XX XX

XX PS

XX Disclosure; Page 126-130; 215pp; English.

XX XX

XX CC

XX This sequence encodes a mouse large tumour suppressor m-lats protein,

XX and has been isolated from a newborn mouse brain phage lambda-ZAP

XX cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene

XX (AAT42117). A homologous mouse sequence has also been isolated

XX (m-lats2, AAT42120). The gene encodes a putative protein-

XX serine/threonine-kinase, and inhibits cell proliferation and plays a

XX crucial role throughout development. Activators or inhibitors of

XX lats function (e.g. an antisense oligonucleotide or dominant-negative

XX proliferative disorders, degenerative disorders, trauma, growth

XX deficiency, etc., and fragments of the gene may be used as diagnostic

XX probes. A lats-inhibitor sequence may be expressed in a transgenic

XX plant or farm animal to confer increased growth and inhibit

XX senescence.

XX XX

XX SQ

XX Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;

Query Match 9.1%; Score 522.8; DB 17; Length 3213;

Best Local Similarity 65.3%; Pred. No. 2.1e-100;

Matches 850; Conservative 4; Mismatches 421; Indels 27; Gaps 5;

QY 2984 AAGGAGAAGGAGGAGGAGGAGGAGTCCGATCAGGAGTACTCGCGGAGGAGGCTTC 3043

DB 1354 AACAAGAAGATCAAGACGACGAGAGTCTCGGATTCAGAGTTACTCCCGACAGGCTTT 1413

QY 3044 AGTTCTTCATGAGGAGGAGGAGGAGGAGTCAATCAAGTGTATCGCCACGAGTAT 3103

Db	2488	AAAGACGGTGCTGATGAGATAAAGGCTCATCATTTTTTTTAAAGACCATCGATTTCTCTAGT	2544		
Qy	4163	GACATCGGAAGCAGAAAGCGCCCTACATACCGGAAATCAAGCACCACCGGACACATCC	4222		
Db	2548	GATCTGAGACGAGCTGCTCTTCATACATCCCTAAATACACGATCCACAGATACATCC	2607		
Qy	4223	AACTTTGATCCCGTGATCCGGAGAAGCTGCGCTCGAATGAC	4264		
Db	2608	AAITTCACCCCTTGTGATCTCTGATAAATGTGGAGCGATGCG	2649		
RESULT 15					
AAZ51506	AAZ51506 standard; DNA; 3213 BP.				
XX	AAZ51506;				
XX	21-JUN-2000	(first entry)			
XX	Mouse Lats (large tumour suppressor) DNA.				
XX	Mouse; Lats; large tumour suppressor; cytostatic; vulnery;				
KW	cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;				
KW	treatment; prevention; screening; cancer; skin; ovarian tumour;				
KW	soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;				
KW	LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;				
KW	displasia; degenerative disorder; growth deficiency; physical trauma;				
KW	hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.				
OS	Mus musculus.				
XX					
Key	Location/Qualifiers				
CDS	1..2889				
FT	/*tag= a				
FT	/product= "Lats protein"				
FT	/partial				
XX	WO200010602-A1.				
PN					
XX	02-MAR-2000.				
XX	18-AUG-1999;	99WO-US19068.			
XX	18-AUG-1998;	98US-0096996.			
PR	18-AUG-1998;	98US-0096997.			
XX	(UYYA) UNIV YALE.				
PA					
XX					
PI	Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;				
PI	Trenchalk GS, Stewart RA;				
XX					
DR	WPI: 2000-246496/21.				
DR	P-PSDB; AA70391.				
XX					
PT	Use of lats proteins, complexes of lats and cdc2 for treating cancer				
PT	that is refractory to treatment by standard chemotherapy and radiation				
PT	therapy, and disorders associated with aberrant levels of cdc2 activity,				
XX					
PS	Claim 44; Fig 13; 134pp; English.				
XX					
CC	The present sequence is a DNA encoding mouse Lats (large tumour				
CC	suppressor) protein which is a cell overproliferation inhibitor and a				
CC	negative regulator of cell cycle-dependent kinase cdc2/cyclin A.				
CC	The present sequence is useful for treating cancer that is refractory				
CC	to standard chemotherapy or radiation therapy such as hyperplasia,				
CC	metaplasia, or dysplasia, and disorders associated with aberrant				
CC	levels of cdc2 activity. Conditions treated by promoting cdc2 function				
CC	include degenerative disorders, growth deficiencies, hypoproliferative				
CC	disorders, physical trauma, lesions, and wounds. An animal model:				
CC	preferably a mouse, in which a lats gene has been disrupted by homologous				
CC	recombination, e.g. a lats knock-out mouse, is used for screening				
CC	compounds that can be used to treat or prevent cancer, particularly				

CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
CC hypogonadotropic hypogonadism. The *lars* DNA is also used in gene therapy.
xx

SQ Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;

Query Match..	9.1%;	Score 522.8;	DB 21;	Length 3213;
Best Local Similarity	65.3%;	Pred No. 2.1e-100;		
Matches 850;	Conservative	4;	Mismatches 421;	Indels 27; Gaps 5;
QY 2984	AGGAGAAGGAGGAGCGCAAGGATTCGCGATCAGCGAGTCTCCGATCCGCGCAAGCCCTTC	3043		
DB				
DB 1354	AACAAGAAAGATGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCCGACAGSCCTTT	1413		
QY 3044	AAAGTTCTTCATCGGAGCAGCACATACAGAACGGTGATCAAGTCGTATCGCGCCAGCGGCACGTAT	3103		
DB				
DB 1414	AAAGTTCTTCATCGGAGCAGCACGTAGAACGCTCTCGAAGTCTCATACAGACGGTCTGCAT	1473		
QY 3104	CCGACGATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCGATCAGACGCCCAATC	3163		
DB				
DB 1474	CGGAAAGACAGCTAGAAAATGAATGATCGGGTTGGATTATCTCAAGATGCCCAGGAT	1533		
QY 3164	GAGATGAGGAAATCCTCAACCAAAGAGAGACACTACATTCGATTAAGCGCGCCCAAG	3223		
DB				
DB 1534	CAATGAGAAAGTCTTTGCCAGAAAGAGTCTAACTATATTCGTCCTTAAAGGCGCTAAA	1593		
QY 3224	ATGGACAAGACATGTTTCGTCAAAACCTGAAGCCCATTTGGAGTGGGTGCTATTTGGCGAGGTA	3283		
DB				
DB 1594	ATGGACAAGCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAGCGTTTGGTGAAGTC	1653		
QY 3284	ACGCTGGTGACCAAAATCGATTACCTCGAACCATTTGTATGGATGAAGAAACCTCGCGGAAA	3343		
DB				
DB 1654	TGCTAGCAAGAAAGTCGGATA---CTAAAGCTTTGTATGCAACAAGACTCTTCGAAAG	1710		
QY 3344	CGGACGTTCTCAAGCGGAATACAGTGGCACACGTGAAGCCGAGAGGGATATCCTCGCG	3403		
DB				
DB 1711	AAAGACGTTCTGCTCCGAAATACAGTGGCTCATGTGAACGGAGAGGGATATCCTAGCA	1770		
QY 3404	GAACCGCAATACTGGGTGGTGAAGTTGTACTACAGCTTCCAGACAAGGATAATCTG	3463		
DB				
DB 1771	GAACCGCAATAAGTGGGTGGTGGCGCTGTACTCTTTCCAGGCAAGGACAACTTG	1830		
QY 3464	TACTTTGTGATGGACTACATACCAGGTGGTGATCTGATGCTGCTCATCAACTGGCG	3523		
DB				
DB 1831	TACTTTGTGATGGACTACATTCCTGGGGGGATATGATGACCTATTAATAGATGGCG	1890		
QY 3524	ATTTTCAGGAGGACATGGCCAGATTCATATCGCGAGGTACCTGCGCGGTGGACAGC	3583		
DB				
DB 1891	ATCTTCTCTGAAATCTGGCCAGCATTCATATGCAAACTTACCTGTGCAGTTGAAAGT	1950		
QY 3584	GTTCAAAAATGGCTTCATTCACAGAGACATCAAGCCTGACAACATATCTCATGATAG	3643		
DB				
DB 1951	GTTCAAAAATGGTTTATTTCATAGAGATTAACCTGTATACATTTGATGATGACCGT	2010		
QY 3644	GACGGACATTAAGCTCACCGACTTTGGCTGTGTGACGGGATTCGGATGGACGCACAC	3703		
DB				
DB 2011	GATGGCCCATTAATAATTGACTGACTTTGGCTGTGCACTGGCTTCAGATGGACACATGAC	2070		
QY 3704	TCCAAGTACTTACCAGGAGAACGCAATCACTCGCCCGAGGACTCGATGGAGCCCTGGGAG	3763		
DB				
DB 2071	TCCAAGTACTACC---AGAGTGGGGATACCCACCGCAAGATAGGATGGATTTCACTAAC	2127		
QY 3764	GAATACTCCGA-----GAACGACCGCAAGCCCAACCGCTGCTGGAGAGCGCA	3808		
DB				
DB 2128	GAATGGGGAGATCTTCCCAATTTGTCGGTGTGGGACAGACTGAACCCACTGGAGCGGAGA	2187		
QY 3809	CGGATCGCGGATCACCAAGAGTCTGCGCCCACTCGCTGGTGGGCAACCCGAACTACATA	3868		
DB				
DB 2188	GCTGCTCGCCACGACACCGGATGTCTAGCCCAATCTCTGGTGGGACTCCCAATATATT	2247		
QY 3869	GTCGCCGAGGCTGGAGAGTAGTGGGTACACGCAAGCTGTGCGACTACTGAGGCGTGGCG	3928		
DB				
DB 2248	GCACCTGAAGTCTACTGCGAACAGGATATACACAGCTGTGTGACTGTGGAGGTGGT	2307		

[illegible]

Search completed: January 16, 2003, 10:36:52
Job time : 895.592 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - 'nucleic search, using sw model

Run On: January 16, 2003, 09:56:02 ; Search time 116.379 Seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1 atctagcagcagcgagca.....gccttgccaaattagttac 5720

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PT05_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5708	99.8	5720	US-09-442-100-1	Sequence 1, Appli
2	5678	9.9	3155	US-09-442-100-7	Sequence 7, Appli
3	557	9.7	1961	US-09-509-902A-15	Sequence 15, Appli
4	536	9.4	1498	US-09-509-902A-6	Sequence 6, Appli
5	522.8	9.1	3213	US-09-442-100-5	Sequence 5, Appli
6	470.6	8.2	3984	US-09-442-100-3	Sequence 3, Appli
7	273.6	4.8	638	US-09-328-111-26	Sequence 26, Appli
8	261.2	4.6	2101	US-08-860-150-1	Sequence 1, Appli
9	261.2	4.6	2101	US-09-338-132-1	Sequence 1, Appli
10	254.8	4.5	2160	US-09-588-256-1	Sequence 1, Appli
11	232.6	4.1	3018	US-08-860-150-6	Sequence 6, Appli
12	232.6	4.1	3018	US-09-338-132-6	Sequence 6, Appli
13	216.6	3.8	1935	US-08-878-989-11	Sequence 11, Appli
14	216.6	3.8	1935	US-09-272-796-11	Sequence 11, Appli
15	164.6	2.9	678	US-09-328-111-66	Sequence 66, Appli
16	135.8	2.4	3182	US-08-484-044-11	Sequence 11, Appli
17	134.6	2.4	2726	US-08-422-699A-12	Sequence 12, Appli
18	134.6	2.4	2726	US-08-422-706B-12	Sequence 12, Appli
19	132.2	2.3	2511	US-08-422-699A-8	Sequence 8, Appli
20	132.2	2.3	2511	US-08-422-706B-8	Sequence 8, Appli
21	126.8	2.2	1599	US-09-256-465-1	Sequence 1, Appli
22	126.8	2.2	1599	US-09-167-322-3	Sequence 3, Appli
23	120.6	2.1	2549	US-09-467-082-3	Sequence 3, Appli
24	116.4	2.0	1732	US-09-430-564-1	Sequence 1, Appli
25	113.6	2.0	3489	US-08-728-323A-1	Sequence 1, Appli
26	113.6	2.0	3489	US-09-298-568-1	Sequence 1, Appli
27	113.6	2.0	32207	US-08-770-379-20	Sequence 20, Appli

c 28	113.6	2.0	32207	4	US-08-757-669A-20	Sequence 20, Appli
c 29	113.6	2.0	32207	4	US-09-230-371A-20	Sequence 20, Appli
30	112.8	2.0	4739	3	US-08-685-871-1	Sequence 1, Appli
31	110.8	1.9	4363	2	US-08-685-576-5	Sequence 5, Appli
32	110.2	1.9	2104	4	US-09-313-930-1	Sequence 1, Appli
33	110	1.9	2610	2	US-09-212-771-1	Sequence 1, Appli
34	110	1.9	2610	3	US-09-091-058-1	Sequence 1, Appli
35	109.2	1.9	5053	2	US-08-685-576-2	Sequence 2, Appli
36	107	1.9	1637	2	US-08-966-316-10	Sequence 10, Appli
37	106.2	1.9	1891	3	US-09-289-466-1	Sequence 1, Appli
38	104.6	1.8	1890	3	US-09-289-466-2	Sequence 2, Appli
39	104.6	1.8	1929	2	US-09-016-000-10	Sequence 10, Appli
40	103.8	1.8	2244	3	US-09-094-714A-48	Sequence 48, Appli
41	103.8	1.8	2245	4	US-09-225-749-24	Sequence 24, Appli
42	103.8	1.8	2599	6	5266464-1	Patent No. 5266464
43	102	1.8	477	4	US-09-135-994-1	Sequence 1, Appli
44	101.4	1.8	397	3	US-09-253-691-3	Sequence 3, Appli
45	99	1.7	2557	4	US-08-464-954A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-442-100-1
; Sequence 1, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111.
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5720 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1103..4402

QY 2101 CACCTCGGGGCTCGCGAGCCCCATAACTGTGTCSTGTGCGCGCGCGCCGCTGGCGAA 2160
DB 2101 CACCTCGGGGCTCGCGAGCCCCATAACTGTGTCSTGTGCGCGCGCGCCGCTGGCGAA 2160
QY 2161 GCCAACAACACAGTCTACCGAGCCAGGAGTCAGCAGCCGATCATCATCAGAGTGTGAA 2220
DB 2161 GCCAACAACACAGTCTACCGAGCCAGGAGTCAGCAGCCGATCATCATCAGAGTGTGAA 2220
QY 2221 GAGCAGCAGAGTCTCAAAAGCCGCTGCTCAAAACAGCAGTGGCGCGCAATGCCATCGAG 2280
DB 2221 GAGCAGCAGAGTCTCAAAAGCCGCTGCTCAAAACAGCAGTGGCGCGCAATGCCATCGAG 2280
QY 2281 TGCCTCGGCCAGCAATCAACAGTCCAGTGTGTCGCGCTCCACCTCTTACCTCTCAGAA 2340
DB 2281 TGCCTCGGCCAGCAATCAACAGTCCAGTGTGTCGCGCTCCACCTCTTACCTCTCAGAA 2340
QY 2341 GTCGCGGCGAGTGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2400
DB 2341 GTCGCGGCGAGTGTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 2400
QY 2401 TCAGCAGCAGCAATCCAAACACCAACGCMACCAACGCGCTTGGTGGGTCTGAAYAG 2460
DB 2401 TCAGCAGCAGCAATCCAAACACCAACGCMACCAACGCGCTTGGTGGGTCTGAAYAG 2460
QY 2461 CAAGCCCAATTCGCTGGAGCCACCTCTCTATGCCAAGAGCATGAGGCGGCGCCAC 2520
DB 2461 CAAGCCCAATTCGCTGGAGCCACCTCTCTATGCCAAGAGCATGAGGCGGCGCCAC 2520
QY 2521 GGTGTARAGCAGCAGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2580
DB 2521 GGTGTARAGCAGCAGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2580
QY 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
DB 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
QY 2641 TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCCAACGGAAA 2700
DB 2641 TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCCAACGGAAA 2700
QY 2701 TCCTGGCGGAGATGCTTCGCGCGCGCGCTTATCAGAGCAACACACACAGGCA 2760
DB 2701 TCCTGGCGGAGATGCTTCGCGCGCGCGCTTATCAGAGCAACACACACAGGCA 2760
QY 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCAACCTGGCGAC 2820
DB 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCAACCTGGCGAC 2820
QY 2821 GACACACCCATTCGCGCTGCGCAATACAAATACAACTCTTCCAAACGCGGCGCAATAG 2880
DB 2821 GACACACCCATTCGCGCTGCGCAATACAAATACAACTCTTCCAAACGCGGCGCAATAG 2880
QY 2881 CTCGGGCGGACGACGAGTCCACGGGACGACGACGACGACGACGACGACGACGACGACG 2940
DB 2881 CTCGGGCGGACGACGAGTCCACGGGACGACGACGACGACGACGACGACGACGACGACG 2940
QY 2941 GATCAAGCAGCGCTCGCCCATCCCGGAGCGCAAGAGATCTCAAGAGAGAGGAGGA 3000
DB 2941 GATCAAGCAGCGCTCGCCCATCCCGGAGCGCAAGAGATCTCAAGAGAGAGGAGGA 3000
QY 3001 GGGCAAGGAGTTCGCATCAGCAGTACTCGCGCGCAAGCTTCAAGTCTTTCATGGAGCA 3060
DB 3001 GGGCAAGGAGTTCGCATCAGCAGTACTCGCGCGCAAGCTTCAAGTCTTTCATGGAGCA 3060
QY 3061 GCACATAGAGACGTGATCAAGTCTGATCGCAGCGCAGCTATCCAGAGATCAGCTGGA 3120
DB 3061 GCACATAGAGACGTGATCAAGTCTGATCGCAGCGCAGCTATCCAGAGATCAGCTGGA 3120
QY 3121 GAAGGAGATGCAACAAGTGGGACTCCCGATCAGACCCCAAAATCGAGATGAGGAAAATGCT 3180
DB 3121 GAAGGAGATGCAACAAGTGGGACTCCCGATCAGACCCCAAAATCGAGATGAGGAAAATGCT 3180
QY 3181 GAACCAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGACAGAGCATGTT 3240

DB 3181 GAACCAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGACAGCATGTT 3240
QY 3241 CGTCAAACTGAAGCCCATTTGGAGTGGTGCAATTTGGCAGAGTAAAGCTGGTGAGCAAAAT 3300
DB 3241 CGTCAAACTGAAGCCCATTTGGAGTGGTGCAATTTGGCAGAGTAAAGCTGGTGAGCAAAAT 3300
QY 3301 CGATACCTCGAACCATTGTTATCGCATGAAAACCTTCGCGAAGCGAGCTTCTCAAGCG 3360
DB 3301 CGATACCTCGAACCATTGTTATCGCATGAAAACCTTCGCGAAGCGAGCTTCTCAAGCG 3360
QY 3361 GAATCAGGTGGCACACGCTGAAGCCGAGAGGATATCTTCGCGAAGCGAGCAATAAATG 3420
DB 3361 GAATCAGGTGGCACACGCTGAAGCCGAGAGGATATCTTCGCGAAGCGAGCAATAAATG 3420
QY 3421 GGTGTGAAGTGTACTACAGCTTCAGGACAAGGATATCTGTTGTGATGGACTA 3480
DB 3421 GGTGTGAAGTGTACTACAGCTTCAGGACAAGGATATCTGTTGTGATGGACTA 3480
QY 3481 CATACAGGTGTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGGAAC 3540
DB 3481 CATACAGGTGTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGGAAC 3540
QY 3541 GGCAGATTCTACATCGCGAGGTCACTGCGCCGCTGGACACGCTTCACAAAATGGGCTT 3600
DB 3541 GGCAGATTCTACATCGCGAGGTCACTGCGCCGCTGGACACGCTTCACAAAATGGGCTT 3600
QY 3601 CATTCACAGAGACATCAAGCCTGACAACATCTCATCTAGTAGGAGGAGACACATAAAGT 3660
DB 3601 CATTCACAGAGACATCAAGCCTGACAACATCTCATCTAGTAGGAGGAGACACATAAAGT 3660
QY 3661 CACCGACTTGGCTGTGACGGGATTCGATGGAGCAGCAACTCGAAGTACTACCAAG 3720
DB 3661 CACCGACTTGGCTGTGACGGGATTCGATGGAGCAGCAACTCGAAGTACTACCAAG 3720
QY 3721 GAACGGCAATCACTCGCGCAGGACTCGATGAGGCCCTGGGAGGAAATCTCCGAGAACGG 3780
DB 3721 GAACGGCAATCACTCGCGCAGGACTCGATGAGGCCCTGGGAGGAAATCTCCGAGAACGG 3780
QY 3781 ACCGAAGCCACCGTGTGAGAGGCGGAGGATGCGGATCAACAAAGAGTCTTGGCCCA 3840
DB 3781 ACCGAAGCCACCGTGTGAGAGGCGGAGGATGCGGATCAACAAAGAGTCTTGGCCCA 3840
QY 3841 CTCGCTGTGGGACCCCGCACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTACAC 3900
DB 3841 CTCGCTGTGGGACCCCGCACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTACAC 3900
QY 3901 GCAGCTGTGCGACTACTGGAGCGTGGGCTCATCTTAYAGATGCTGGTGGGTACGCC 3960
DB 3901 GCAGCTGTGCGACTACTGGAGCGTGGGCTCATCTTAYAGATGCTGGTGGGTACGCC 3960
QY 3961 GCCCTTTCGCGCAACAGTCCGCTGGAACGCAACAAAGGTCATCACTGGGAGAAAC 4020
DB 3961 GCCCTTTCGCGCAACAGTCCGCTGGAACGCAACAAAGGTCATCACTGGGAGAAAC 4020
QY 4021 SCTGCATATTCGCGCGAGCGGAGTTCATCCCGGAGGCTACGAGCTTGTATAGAGGCT 4080
DB 4021 SCTGCATATTCGCGCGAGCGGAGTTCATCCCGGAGGCTACGAGCTTGTATAGAGGCT 4080
QY 4081 CTGTCGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACGAGGTCAAGAGCCAGACTT 4140
DB 4081 CTGTCGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACGAGGTCAAGAGCCAGACTT 4140
QY 4141 CTTCAAGGCACTGACTTTCGCGGATCGGAGCAGAAAGCGCCCTACATACCCGAAT 4200
DB 4141 CTTCAAGGCACTGACTTTCGCGGATCGGAGCAGAAAGCGCCCTACATACCCGAAT 4200
QY 4201 CAAGCACCRCAGGACACATCAACTTTCATCCCGTGTATCCGAGAGAGCTCGCTCGAA 4260
DB 4201 CAAGCACCRCAGGACACATCAACTTTCATCCCGTGTATCCGAGAGAGCTCGCTCGAA 4260
QY 4261 TGACTCCACCATGAGCAGCGGATGATGTGACCAAGATGACCGCACCTTCCACGGCTT 4320

Fri Jan 17 11:17:10 2003

Db 4261 TGACTCCACCATGAGCGGGCGATGATGTCGACCAAGATGACCGACATTTCCACGGCTT 4320
Qy 4321 TTTTCGAATTTACCTTCCTCGCTGCTTCTTCGACGACAGACGCGCGGATATGACGGACGA 4380
Db 4321 TTTTCGAATTTACCTTCCTCGCTGCTTCTTCGACGACAGACGCGCGGATATGACGGACGA 4380
Qy 4381 TCAGGCGCGGTTTACGTCFAGAAATGGATGCTCTCCATGTGCGCAACACACACCCCGC 4440
Db 4381 TCAGGCGCGGTTTACGTCFAGAAATGGATGCTCTCCATGTGCGCAACACACACCCCGC 4440
Qy 4441 CCCCGAATCATTTAGTCAAAATAGTCACAAAAGGGGATAGAAACCAATGAGTGGGCTT 4500
Db 4441 CCCCGAATCATTTAGTCAAAATAGTCACAAAAGGGGATAGAAACCAATGAGTGGGCTT 4500
Qy 4501 GCATTTGTAAGGAGCGTGGCTATAGAAATGAAACTATCTATATACATATATAAATATA 4560
Db 4501 GCATTTGTAAGGAGCGTGGCTATAGAAATGAAACTATCTATATACATATATAAATATA 4560
Qy 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATTAATATACATATATTCATAT 4620
Db 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATTAATATACATATATTCATAT 4620
Qy 4621 ATATATATATATATGCGGTAGGCGTACGTAAGTGAATTAATATAAAGCGGAGCGTAG 4680
Db 4621 ATATATATATATATGCGGTAGGCGTACGTAAGTGAATTAATATAAAGCGGAGCGTAG 4680
Qy 4681 AGATGAAACGAGAGCGGAGCTCGAGCCTTCGACCTTTAACTGAACATAGTATATCCTT 4740
Db 4681 AGATGAAACGAGAGCGGAGCTCGAGCCTTCGACCTTTAACTGAACATAGTATATCCTT 4740
Qy 4741 GTGCACTACTACTCCACAAATATATATTTTAAATTTAGAAATTCAAAAGGGGCCA 4800
Db 4741 GTGCACTACTACTCCACAAATATATATTTTAAATTTAGAAATTCAAAAGGGGCCA 4800
Qy 4801 ACTGGAATCGAACTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCC 4860
Db 4801 ACTGGAATCGAACTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCC 4860
Qy 4861 TTAACATAAAGAGAGCGCAATTTACCAACCACTTCACTCTCTCTCTCTCTCTCTCTCC 4920
Db 4861 TTAACATAAAGAGAGCGCAATTTACCAACCACTTCACTCTCTCTCTCTCTCTCTCTCC 4920
Qy 4921 GATCGTGGCGGATTCGAACCTCAGCAGGCTGGTTCATCCGCGCATCCCATTTKACTTCC 4980
Db 4921 GATCGTGGCGGATTCGAACCTCAGCAGGCTGGTTCATCCGCGCATCCCATTTKACTTCC 4980
Qy 4981 CATTCAGAAATGAGATTCGAGGTGCGATGAGACGACGAGACGACCAAGTCGCAC 5040
Db 4981 CATTCAGAAATGAGATTCGAGGTGCGATGAGACGACGAGACGACCAAGTCGCAC 5040
Qy 5041 GCGAGCGATATAGCGGCTTTATAGCCCTTAATCTAAATCTAACTGGGAGAACAGGACC 5100
Db 5041 GCGAGCGATATAGCGGCTTTATAGCCCTTAATCTAAATCTAACTGGGAGAACAGGACC 5100
Qy 5101 YATGTATGCTCTATFCCAAATTCGCTATCATCTGCTCTTCAWCTGWTGACACCCAC 5160
Db 5101 YATGTATGCTCTATFCCAAATTCGCTATCATCTGCTCTTCAWCTGWTGACACCCAC 5160
Qy 5161 CCCCCCTCCCTCCATCCAAAGAACAACTTACAGCTAGCCTATGTGAAAAGCTAGCAAT 5220
Db 5161 CCCCCCTCCCTCCATCCAAAGAACAACTTACAGCTAGCCTATGTGAAAAGCTAGCAAT 5220
Qy 5221 GTTAGACCAACTTGTGAATGCCAAATGAAATGTTTACGCCACGAGGAAACGGGG 5280
Db 5221 GTTAGACCAACTTGTGAATGCCAAATGAAATGTTTACGCCACGAGGAAACGGGG 5280
Qy 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
Db 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
Qy 5341 AAACAGTACGAGAAATTTGTAATCTCTTATGTAATGTAATGTAATGTAATGTAATGTA 5400
Db 5341 AAACAGTACGAGAAATTTGTAATCTCTTATGTAATGTAATGTAATGTAATGTAATGTA 5400

Qy 5401 AATCTATGCTAGAGCTTGTGTAGCGCCCTAAGATGTTTTTTTAGTTTATAGACCCTAACCG 5460
Db 5401 AATCTATGCTAGAGCTTGTGTAGCGCCCTAAGATGTTTTTTTAGTTTATAGACCCTAACCG 5460
Qy 5461 TAACTCTAGTTTAAATTCCTAACACATAAGCGAGAGTACAGTACATTTGTTTGTGTC 5520
Db 5461 TAACTCTAGTTTAAATTCCTAACACATAAGCGAGAGTACAGTACATTTGTTTGTGTC 5520
Qy 5521 GTAGTTCGTTGGAAATGCTTAAACGGAAGAGATTTGTTTCTCTTTAATTAGCTTCA 5580
Db 5521 GTAGTTCGTTGGAAATGCTTAAACGGAAGAGATTTGTTTCTCTTTAATTAGCTTCA 5580
Qy 5581 GTTTCATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
Db 5581 GTTTCATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
Qy 5641 TGGAGCCTATTTAAATGTGAGATCGAGCTAAATGAAGAAATACAAACAACTCTGTGT 5700
Db 5641 TGGAGCCTATTTAAATGTGAGATCGAGCTAAATGAAGAAATACAAACAACTCTGTGT 5700
Qy 5701 GCCTTGGCCAAATAGTTTAC 5720
Db 5701 GCCTTGGCCAAATAGTTTAC 5720

RESULT 2
US-09-442-100-7
Sequence 7, Application US/09442100
Patent No. 6359193
GENERAL INFORMATION:
APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Welyi
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirostock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

FEATURE: CDS
NAME/KEY: 1..2943
LOCATION: 1..2943
US-09-442-100-7

Query Match
Best Local Similarity 64.3%; Score 567.8; DB 4; Length 3155;
Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCGGATAGCTCGGGGCGGACGAGTCCACGCGGACGCGCTCTCTCGTCGAC 2928
DB 1329 GGGCGGATAGCTCGGGGCGGACGAGTCCACGCGGACGCGCTCTCTCGTCGAC 1388

QY 2929 CAGCTCCAAAGAGATCAAGACGCGCTCGGCCATCCCGGAGCGCAAGAGATCTCCAAGGA 2988
DB 1389 CAGAGACAAAGACGATTCAGACCTCCCGGCTGCTCGGCAAGATAGC-----1440

QY 2989 GAAGGAGGAGGCGCAAGAGATTCGCGATCAGGCGAGTACTCGCCGCAAGCCTTCAAGTT 3048
DB 1441 -AGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTACTCCCTTTATGCGCTTCAATT 1499

QY 3049 CTTATGGAGGACGACATAGAGACGATGATCAAGTGGGACTCGCCGATCAGACCCAAATCGAGAT 3108
DB 1500 CTTATGGAGGACGACGATGAGAGTCTCAAAACCTACCAGCAAGGTCAGCCCGAG 1559

QY 3109 GAATCAGCTGGGAGAGGAGATGCACAAAGTGGGACTCGCCGATCAGACCCAAATCGAGAT 3168
DB 1560 GCTACAGCTGGGAGAGGAGATGCACAAAGTGGGACTCGCCGATCAGACCCAAATCGAGAT 1619

QY 3169 GAGGAAATGCTGAACAAAGGAGAGCAACTACATTCGATTGAAGCGCGCCAAAGATGA 3228
DB 1620 GAGGAAATGCTGAACAAAGGAGAGTCTAACTACAAAGGCTGAAGAGGCGCAAGATGA 1679

QY 3229 CAAGACATGCTCGTCAAACTGAGCCCAATGGAGTGGGTGCTATTTGGCGAGGTACGCT 3288
DB 1680 CAAGTCCATGTTGTGAAAATCAAGACTCTAGGCAATCGGTGCTTTGGGGAAGTGTGCT 1739

QY 3289 GGTGAGCAAAATCGATACCTCGAACCATTTGATGGGATGAAACCTCGGAAAGCGGA 3348
DB 1740 CGCTTGTAGCTGA---CACTCAGCTGTGTACGCCATGAAGCTCTCAGGAAGAGA 1796

QY 3349 CGTTCTCAAGCGAATCAGTGGGACGACGTCGAGGCGGAGGATATCTCGCGGAAGC 3408
DB 1797 TGTCTGAAACCGAATCAAGTGGCCATGTCAAGGCTGAGAGGACATCTGTGCTGAAGC 1856

QY 3409 CGACATACTGGTGGTCAAGTGTGCTACAGCTTCCAGGACAGGATTAATCTGTACTT 3468
DB 1857 AGCAATAGTGGTGGTCAAGTGTGCTACCTCTCCAGGACAGGACGCTGTACTT 1916

QY 3469 TGTGATGAGTACATACGAGTGGTGTGATGCTGCTCATCAAACTGGGCTTTT 3528
DB 1917 TGTGATGAGTACATACGAGTGGGATATGATGAGCTGTGATCAGGATGGAGTCTT 1976

QY 3529 CGAGGAGAACTGGCGAGATTTACATCGCCGAGGTACCTGGCGGCTGGACAGCGTTCA 3588
DB 1977 CCGTGAACCTGGCGGCTTTACATTCGAGATTGACCTGGCCATGAAAGTGTCCA 2036

QY 3589 CAAATGGCTTTCATTCAGAGACATCAAGCTGACACATACCTATCGATAGGAGCGG 3648
DB 2037 CAAGTGGGCTTTATCCACGGGACATCAAGCTGACACATACCTATCGATAGGAGCGG 2096

QY 3649 ACATATAAGCTCACCGGCTTTGGGCTGTGCGAGGATTCGATGGACGCAACCTCGAA 3708
DB 2097 TCATATTAGCTGACAGATTTGGGCTCTGCTGATGAGTTCAGTGGACATCAATTCGAA 2156

QY 3709 GTACTACGAGGAGCGCAATCACTCGCCGAGGACTCGATGGAGCCCTGGGAGGAAT- 3767
DB 2157 GTACTACC---AGAAAGGAACACATGAGACGAGACATGGAGCCCGGTGACCTCTG 2213

QY 3768 -----ACTCCGAGACGGACCGAGCCGAGCCCGTGTGAGAGGGGACGGAT 3813
DB 2214 GGACGATGTTTCCAACTGTGCTGTGGAGACAGGTTAAAGACCTTGAAGACGAGGCGCA 2273

QY 3814 GGGCGATCACCAGAGAGTCTGGCCACATCTGCTGGTGGGACCCCGAACTACATAGCTCC 3873
DB 2274 GAAGCAGCAGAGAGTCTGGCAGATTTCTTGTGCGGACACCAATATACATCGCTCC 2333

QY 3874 CGAGTGTGGAGAGAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGGCGTCTAT 3933
DB 2334 GGAGTGTCTTCTCGGCAAGAGGTACACGAGCTCTGTGACTGGTGGAGCGTGGTGTGAT 2393

QY 3934 CCTTATAGAGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 3993
DB 2394 TCTTATAGAGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2453

QY 3994 ACAAAGGTTCATCACTGGGAGAAACCTGCATATTCGCGCGAGCGGAGTTATCCGG 4053
DB 2454 GCTGAAGTTCATCACTGGGAGAGACGCTGCATATCCCTAGCAGGTGAGCTCAGCG 2513

QY 4054 CGAGGCTACGACTTGTATAAGGAGGCTGTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4110
DB 2514 TGAGGCGCGAGACCTCATCAGAAAGCTGTGCTGGCGGCTGACTGCGGCTGGGAGGA 2573

QY 4111 GAGCGTGGAGAGGTCAGAGGACGACCTTCTTCAAGGCGATCGACTTTGC---GGACAT 4167
DB 2574 TGGGCGAGTACCTCAAGGACACCCGCTTCTCAACACCATCGACTTTTCCCGTGACAT 2633

QY 4168 GCGGAGCAGAAAGCGCCCTACATACCGGAAATCAAGCACCCACGACACATCCAACTT 4227
DB 2634 CGAAAGCAGGCTGACCTTACCTCCACCATCAGCCACCCCTGGACACCTCCATTT 2693

QY 4228 TGATCCGCTGGATCGGAGAGGTCGGCTGCTGATGATGATGATGATGATGATGATGATGAT 4287
DB 2694 TGACCCGCTGGATGAAGAAAGCCCTGGCAGCGAGGCGGAGAGAGCGCCAAAGCCCTG 2753

QY 4288 TGTGCGACAGATGACCGGACCTTTC-----CAGGCTTTTTCGAATTTACCTT 4335
DB 2754 GGACAGCTGGGCTTCCCGCAGCAGCAAGCATCCAGACGCGCTTCTATGAGTTACCTT 2813

QY 4336 CCGTGTCTTCTCGAGCAAA 4356
DB 2814 CCGAGGTTCTTCGATGACAA 2834

RESULT 3
US-09-509-902A-15
; Sequence 15, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:

; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1961
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-15

Query Match
Best Local Similarity 64.1%; Score 557; DB 4; Length 1961;
Matches 963; Conservative 4; Mismatches 494; Indels 42; Gaps 7;

QY 2884 GGGCGGACAGAGTCCACCGGACCACTCCCTCTGTCGACGCTGCAAGAGAT 2943
DB 158 GGGCGGACAGAGTCCACCGGACCACTCCCTCTGTCGACGCTGCAAGAGAT 217

QY 2944 CAAGCAGGCTCGCCCTATCCCGGAGCGCAAGAGATCTCCAAAGGAGAGGAGCGG 3003
DB 218 GATTGAGACCTCTCCGCTTCCCGTCCGCAAAAGAGC-----AGAGACGAGAGAA 268

Qy	3004	CAAGGAGTTC	CCGATCAGG	CAGTACTCG	CGCGCAAGC	CTTCAAGTTC	TCTTCATGG	AGCACA	3063
Db	269	GAGAGATC	ACGATCAAG	AGCTACTCG	CCATACGCC	TTTAAAGTTC	TCTTCATGG	AGCACA	328
Qy	3064	CATAGAGA	CGTGATCA	AGTCTATCG	CGCGCACGT	ATCGCAAG	AATCAGCTG	GGAGAA	3123
Db	329	CGTGGAGA	TGTCAATA	CAACCTAC	CAGCAGA	AGGTTAA	CCGGAGGT	GCAAGCTG	368
Qy	3124	GGAGATGC	ACAAAGT	GGGACTCG	CGCATCAG	ACCACAA	ATCAGATG	AGGAAATG	3183
Db	389	AGAAATGG	CCAAAGCT	GGACTCTG	TGAAGCT	GACAGAG	CAGATGCG	GAAGATC	448
Qy	3184	CCAAAGGA	GAGCAACT	ACATTCG	ATGGAAG	CGCCCAAG	ATGACAGAG	CATGTT	3243
Db	449	CCAGAAGA	GTCTAATT	ACAACAG	GTAAAG	AGGGCCA	AGATGACA	AGTCTAT	508
Qy	3244	CAAACTGA	AGCCATT	GGAATGG	AGTGGTGC	ATTGGCG	AGGTAC	GCTGTG	3303
Db	509	CAAGATCA	AAACCC	TGGGATCG	GTGGCCCT	TTGGAGA	AGTGTGC	TTTGTAA	568
Qy	3304	TACCTCGA	ACCAAT	TGTTATG	CGATGA	AAACCC	CTCGGAA	AGCGAC	3363
Db	569	---CACTC	ACGCCCT	GTACGCC	ATGAAG	ACCCTA	AGGAAAG	GAATGCT	625
Qy	3364	TCAGTGGC	ACAGTGA	AGCGCG	GAGGGAT	ATCCTCG	CGAAGC	GCACAA	3423
Db	626	TCAGTGGC	CCACGT	CAAGCCG	GAGGGAC	ATCCTGC	CGGAGC	AGACAAT	685
Qy	3424	GGTGAAGT	GTACTACA	GCCTTCC	AGGAC	AGGATA	ATCTGCT	TTGTGAT	3483
Db	686	GGTCAAA	CTACTACT	CTCTTCC	AAAGCA	AAAGAC	GCCTGTACT	TTGTGAT	745
Qy	3484	ACCAGTGT	GATGTG	CTGCTCA	TCAACT	GGGCA	TTTTCG	AGGAGAA	3543
Db	746	CCCTGGT	GGGACAT	GATGAG	CCCTGT	GATCCG	GATGG	AGGTCT	805
Qy	3544	CAGATTCT	ACATCG	CGAGGT	TCACCT	CGCGT	GGACAG	CGTTCA	3603
Db	806	CCGGTCT	TACATCG	ACAGCT	GTACTT	GGCCATT	GAGAGT	TCCAC	865
Qy	3604	TCACAGAC	ACATCA	AGCCTG	ACAACT	ATCTG	ATGAGG	ACGACAT	3663
Db	866	CCACGGAC	ATCAAG	CTGTAT	ACATTT	TGATAG	ATCTGG	ATGATTA	925
Qy	3664	CGACTTTG	CGCTGC	ACCGGATT	TCGATG	GGACG	CACAACT	CGAAGT	3722
Db	926	AGATTTG	CGGCTCT	GCATCG	GTTCAG	TGGACT	CACAAT	TCCAAT	985
Qy	3723	-----AC	GGCAAT	CTCTG	CGCCAG	GCATG	ATGAG	CGCCTG	3774
Db	986	GAGCCAT	GTGAG	ACAG	CACAG	CACTATG	AGGCC	ACCTCT	1045
Qy	3775	---GAAC	CGGAC	CGCCAC	CGTGTG	GAGAG	CGCCAG	CGATG	3831
Db	1046	TCGGTGT	GGGACAG	CTGNAG	ACCCTAG	ACAGAG	GGCGG	GAAGAG	1105
Qy	3832	CTTGGCC	CACTCG	TGTGG	GCACCC	GAACCTAC	ATAGCT	CCGAGGT	3891
Db	1106	CTTGGCA	CAATTC	ACTGTG	GGGACT	CCAAACT	ATACAT	CGCAC	1165
Qy	3892	TGGGTAC	ACGAC	GCCTGT	GCGACT	ACTGG	AGCTGG	CGCTCT	3951
Db	1166	AGGGTAC	CTCACT	CTGTG	ACTGG	TGGAGT	TTGGAGT	GATTC	1225
Qy	3952	GGGTAC	AGCCG	CCCTTT	CTTGG	CCAA	CAGTCC	CGTGG	4011
Db	1226	GGGGC	ACGCG	CCCTTT	TTTGG	CACT	TCCAC	ACAGAA	1285
Qy	4012	GGAGAA	AC	CTGCA	TATTTCC	CCCGC	CGCG	AGTTAT	4071
Db	1286	GGAGAA	CA	CGTCC	ACATTTCC	AGCC	CGAG	TGAAGCT	1345

```

QY 4072  AAGGAGGCTCTGCTCGCTCGCTGACAAGCGGCTGGGCAAGAGCGTGGACGAGG---TCAA 4128
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1346  CACCAAGCTGTGCTGCTCCGCAGACACCACCGCTGGGCGGGAATGGGGCGATGACCTGAA 1405
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4129  GAGGCACGACTCTTCAAGGCGATCGACTTTGC---GGACATCGGGAAGCAGAAAGCGCC 4185
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1406  GGCCACACCCCTTCTTCAGCGCAATTGACTTCTCCAGTGACATCCGGAAGCATCCAGCCCC 1465
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4186  CTACATACCGGAATAACAAGACACCCACGACGACACATCCAATTTTGATCCCGTGGATCCGGA 4245
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1466  CTAAGTCTCCACCATCAGCACACCCATGGACACTCGAATTTTCGACCCCGTAGATGAAGA 1525
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4246  GAAGTCGCGCTCGAATGACTCTACCATGAGCAGCGGCGATGATGTC-----GA 4293
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1526  AAGCCCTTTGGAACGATGCCAGGAAGTAGCACCAAGGCTGGGACACACTCACCTCGCC 1585
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4294  CCAGAAATCAGCGCACYTTCACGGCTTTTCGAAATTTACTCTTCGCTTCGTCGACGA 4353
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1586  CAATTAACAAGCATCTCGAGCAGCAATTTACGAATTCACGCTTCCGAAGGTTCTTTGATGA 1645
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 4354  CAA 4356
Db      |||
Db 1646  CAA 1648

RESULT 4
US-09-509-902A-6
; Sequence 6, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US.
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-6

Query Match          9.4%; Score 536; DB 4; Length 1498;
Best Local Similarity 65.5%; Pred. No. 1.8e-118;
Matches 884; Conservative 3; Mismatches 433; Indels 30; Gaps 6;

```

Query Match	9.4%	Score 536;	DB 4;	Length 1498;
Best Local Similarity	65.5%;	Pred.	No. 1.8e-118;	
Matches 884;	Conservative	3;	Mismatches 433;	Indels 30; Gaps 6;
Qy	2884	GGCGGCACGAGGATCCACGCCACCACCGCTCCTCGTCACCAAGTGCAGAAGAT	2943	
Db	160	GGCGGCGACAAGCGCGCAAAAGCGCAAGGGGACAAAAGCGGAAGGATAAAAAGCA	219	
Qy	2944	CAAGCACGCTCGCCCATCTCCGGAGCGCAAGAAGATCTCCAAGGAGAGGAGGAGCGC	3003	
Db	220	GATTCAAGACTCTCCGTTCGGTCGCGAANAACAGC-----AGAGCGAAGAGAA	270	
Qy	3004	CAAGGAGTTCGCATCAGGCAGTAGTACTCGCGCAAGCCTTCAAGTTCTTCATGGAGCAGCA	3063	
Db	271	GAGAGAGTCACGATCAAGAGCTACTCGCCATAGCGCTTTTAAGTTCTTCATGGAGCAGCA	330	
Qy	3064	CATAGAGAAGTGATCAAGTCGATCGCCAGCGCAGCTATCGCAAGAANTCAGCTGGAGAA	3123	
Db	331	CGTGGAGAACTGATCAAAAACCTACCAGCAGAAGTTAACCGGAGGCTCGACGCTGGAGCA	390	
Qy	3124	GGAGATGCACAAAGTGGAGCTGCCGATCAGACCCCAATTCGAGATCAGGAAAAATCGTAA	3183	
Db	391	AGAAAATGGCCAAAAGTGGACTCTGTGAAGCTGAGCAGGAGCAGATCGCGAAGATCCTCTA	450	
Qy	3184	CCAAAAGAGAGCAACTACATTCCATTGAAGCGCCCAAGATGGACAGAGCATGTTCGT	3243	
Db	451	CCAGAAAGAGTCAATTACAACAGGTTTAAAGGGGCGCAAGATGGACAAGCTATCTTTGT	510	

RESULT 6
US-09-442-100-3
; Sequence 3, Application US/09442100
: Patent No 6359193

RESULT 6
US-09-442-100-3
; Sequence 3, Application US/09442100
: Patent No 6359193

Db 2388 TCTAGCAAGAAAGTAGATA---CTAAGGCTTTGTATGCAACAAAACCTCTCGAAG 2444
 QY 3344 GCGGACGTTCTCAAGCGGAATCAGGTGGCACACGTCGAGCGGAGAGGATATCTCGCG 3403
 Db 2445 AAGATGTTCTTCTCGAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTCGGT 2504
 QY 3404 GAAGCGCAATCAATCAAGTGGTGTGAAGTTGTACTACAGCTTCCAGGCAAGGATAATCTG 3463
 Db 2505 GAAGCTGACAATGAATGGTAGTTCGTCTATATTATTCATCCAGATAAGACAAATTA 2564
 QY 3464 TACTTTGTGATGAGTACATACAGGTGGTGTGATCTGCTGCTCATCAAACTGGGC 3523
 Db 2565 TACTTTGTGATGAGTACATCTCTGGGTGATGATGATGAGCTTATTAATGAATGGC 2624
 QY 3524 ATTTTCGAGGAGGAATGCCAGATCTACATCGCGGAGTCACTCGCCCTGGACAGC 3583
 Db 2625 ATCTTTCGAGAAGTCTGGCAGCTTCTACATAGCAGAACTTACCTGTGCACTTGAAGT 2684
 QY 3584 GTTCACAAATGGGCTTCATTCACAGACATCAAGCCCTGACAACATCACTCATCATAGG 3643
 Db 2685 GTTCATAAATGGGTTTATTCATAGAGATATTAACCTGATAATATTGATGATCT 2744
 QY 3644 GAGGACACATTAAGCTCACCGACTTGGCCCTGTGCGACGGATTCGGATGGACGCAAC 3703
 Db 2745 GATGTCATATTAATTAAGTCACTGACTTGGCTCTGCACTGGCTTCAGATGGACACGAT 2804
 QY 3704 TCGAAGTACTACAGGAGGCAATCACTCGCCGAGGACTCGATGAGCCCTGGGAG 3763
 Db 2805 TCTAAGTACTATC---AGAGTGTGACCATCCAGGCAAGATAGATGATTTAGTAAT 2861
 QY 3764 GAATACTCCGAGAAC-----GGACCGAAGCCCAAGCTGTGTGAGAGGCGA 3808
 Db 2862 GAATGGGGATCCCTCAAGCTGTCGATGTGAGACAGACTGAAGCCATTAGCGGAGA 2921
 QY 3809 CGGATGGGGATCACCAAGAGTCTCGCCCACTCGCTGGTGGGACCCCGCAACTACATA 3868
 Db 2922 GCTGACGCGCAGCAGCAGGATGTAGCACATCTTGGTTGGGACTCCCAATTAAT 2981
 QY 3869 GCTCCCGAGGCTCTGAGAGAGTGGGTACACGAGCTGTGCGACTACTGAGGCGTGGC 3928
 Db 2982 GCACCTGAAGTGTCTGACGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGT 3041
 QY 3929 GTCACTCCYATYAGATGCTGGTGGTCAAGCGCCCTTCTGCGCAACAGTCCGCTGGAA 3988
 Db 3042 GTTATCTTTTGAATGTTGGTGGGACAACTCTTCTTGGCAACACACCATAGAA 3101
 QY 3989 AGCAACAAAGGTCATCAACTGGGAGAAACSCGTGCATATTCGCGCGAGCCGAGTTA 4048
 Db 3102 ACACAAATGAAGGTTATCAACTGGCAACATCTCTACATTTCCACCAAGCTAAACTC 3161
 QY 4049 TCCCGGAGGCTACGAGTGTATAGGAGGCTCTGTGCGCTGACAGCGGCTGGC 4108
 Db 3162 AGCTCTGAAGCTCTGTGATCTTATTAATTAACCTTCCGAGGAGCCGAGATCGCTTAGC 3221
 QY 4109 AAGA---CGCTGGAGGAGTCAAGACGACGACTTCTTCAAGGCGATCGACTTTCG---G 4162
 Db 3222 AGAATGCTGTGATGAATAAAGCTCATCCATTTTAAACAAATGACTTCTCCAGT 3281
 QY 4163 GACATCGGAAGAGAAAGCGCCCTACATACCGGAAATCAAGCAACCCACGACATCC 4222
 Db 3282 GACCTGAGACAGCAGCTGCTTTCATACATTCCTAAATACACACCCACAGATACATCA 3341
 QY 4223 AACTTTGATCCGCTGGATCCGAGAGGCTGGCTCGAATGA 4263
 Db 3342 AATTTTGAATCTGTTGATCTCTGATATAATTTATGGAGTGA 3382

RESULT 7

US-09-328-111-26/c
 ; Sequence 26, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:

APPLICANT: Endege, Wilson O.
 APPLICANT: Steinmann, Kathleen E.
 APPLICANT: Astle, Jon H.
 APPLICANT: Burgess, Christopher C.
 APPLICANT: Bushnell, Steven E.
 APPLICANT: Carroll III, Eddie
 APPLICANT: Catino, Theodore J.
 APPLICANT: Dertli, Adnan
 APPLICANT: Ford, Donna M.
 APPLICANT: Lewis, Marcia E.
 APPLICANT: Monahan, John E.
 APPLICANT: Schlegel, Robert
 TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 TITLE OF INVENTION: PRODUCTS
 FILE REFERENCE: CCD-257 (US)
 CURRENT APPLICATION NUMBER: US/09/328,111
 CURRENT FILING DATE: 1999-06-08
 EARLIER APPLICATION NUMBER: US 60/088,801
 EARLIER FILING DATE: 1998-06-10
 NUMBER OF SEQ ID NOS: 850
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 26
 LENGTH: 638
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-09-328-111-26

Query Match 4.8%; Score 273.6; DB 4; Length 638;

Best Local Similarity 66.4%; Pred. No. 6.4e-56;

Matches 425; Conservative 0; Mismatches 209; Indels 6; Gaps 2;

QY 3074 GTGATCAAGTCTGATCGCCAGCGCACGTATCGCAAGAATCAGCTGGAGAGGAGATGCAC 3133
 Db 637 GTACTCAATCTCATCAGCAGCGCTCTACATCGTAAAAACAATTAGAGATGAATGATG 578
 QY 3134 AAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGAGAAATGCTGAACCAAAAGGAG 3193
 Db 577 CGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 518
 QY 3194 ACCAACTACATTCGATTGAAGCGCCCAAGATGGCAAGAGATGTTCTCTCAACTGAG 3253
 Db 517 TCTAATACATCCGCTTTAAAGGGCTAAAATGGACAGTCTATGTTCTGAAGATAAAG 458
 QY 3254 CCATTTGAGTGGGTGCTATTTGGCGAGGTAAACGCTGGTGGCAAAAATCGATACCTCGAAC 3313
 Db 457 ACATAGGATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATA---CTAAG 401
 QY 3314 CATTTGTATGCGATGAAACCCCTGCGGAAAGCGGAGCTTCTCAAGCGGAATCAGGTGGCA 3373
 Db 400 GCTTTGTATGCAACAAAACCTCTCGAAATAAAGATGTTCTTCTCGAAATCAAGTCGCT 341
 QY 3374 CAGCTGAAGCGGAGAGGATATCTCGCGGAGCGGACAAATAACTGGGTGGTGAAGTTG 3433
 Db 340 CATGTTAAGGCTGAGAGAGATATCTCGGCTGAAGCTGACAATCAATGGGTAGTTCGTCTA 281
 QY 3434 TACTACAGCTTCCAGGACAAAGGATATCTGTTGCTGATGAGTACATACAGGTGCT 3493
 Db 280 TATTATTCTCCAGATAGGACAAATTTATCTGTTGTAATGACTACATTTCTCTGGGGT 221
 QY 3494 GATCTGATGCTGCTCATCAAACTGGGCATTTTTCGAGGAGGAACTGGCCAGATTTCTAC 3553
 Db 220 GATATGATGAGCCTATTAAATAGAAATGGGCATCTTTCCAGAAAGTCTGCGCAGGATCTAC 161
 QY 3554 ATCGCGGAGGTCACCTGCGCGCTGGACAGCG---TTCACAAAATGGGCTTCATTACAGA 3610
 Db 160 ATAGCAGAACTTACCTGTGCAAGTGTTCCTTTAAAGTGGGTTTATTTATCATAGA 101
 QY 3611 GACATCAAGCTGACAAACATCTCATAGGAGGAGGACACATAAAGCTCACCAGCTTT 3670
 Db 100 GATATTAAACCTGATAATATTTGATGATCTGATGATGATTAATTAATGACTGACTTT 41
 QY 3671 GGCCTGTGCACGGGATTCGGATGGACGCAACAACTCGAAGT 3710
 Db 3671 GGCCTGTGCACGGGATTCGGATGGACGCAACAACTCGAAGT 3710

Db 40 GGCCTCTGCACTGGCTTCAGATGACACACGATTTAAAGT 1

RESULT 8

US-08-860-150-1

; Sequence 1, Application US/08860150B

; Patent No. 5981205

; GENERAL INFORMATION:

; APPLICANT: Hemmings, Brian A.

; APPLICANT: Millward, Thomas A.

; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases

; FILE REFERENCE: 4-20265/A/PCT

; CURRENT APPLICATION NUMBER: US/08/860,150B

; EARLIER FILING DATE: 1997-06-19

; EARLIER APPLICATION NUMBER: PCT/EP95/05052

; EARLIER FILING DATE: 1995-12-20

; EARLIER APPLICATION NUMBER: 94810746.1

; EARLIER FILING DATE: 1994-12-22

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2101

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (132)..(1499)

US-08-860-150-1

Query Match 4.68; Score 261.2; DB 2; Length 2101;

Best Local Similarity 55.08; Pred. No. 1.1e-52;

Matches 581; Conservative 3; Mismatches 451; Indels 21; Gaps 3;

QY 3043 CAAGTCTTCATGAGCAGACATAGAGACGTCATCAAGTCGTATGCCAGCGCAGTA 3102

Db 212 CAAGGTGACGTTGGAGAACTACTACAGCAACCTGGTGACGAGTATGGCGAGGAAAGCA 271

QY 3103 TCGCAAGAATCAGCTGGAGAGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAAT 3162

Db 272 GCGCTCGCAAGCTGGAGCTCAGCTGAGGAGGAGAGCTTGTGGAGCGGAGCGCCA 331

QY 3163 CGAGATGAGGAAATGCTGAACCAAGAGAGAGCACTACATCGATGTAAGCGCGCAA 3222

Db 332 GGAGAGCGCTGTCAGCATGCCAGAGAGAGAGGAGTATCTCCGGCTGAAGCGCATTCG 391

QY 3223 GATGCAAGAGCATGTTCTGCAAACTGAAGCCATTTGGAGTGGTGCATTGGCGAGGT 3282

Db 392 CTTGGTGTGAGGACTTTGAGGCCCTCAAGTCATCGGAGCGGCGCTTGGTGAAGT 451

QY 3283 AACGCTGGTGAAGCAAAATCGATACCTCGAACCACTTTGTATGCGATGAAACCCCTGCGAA 3342

Db 452 GCGTTTGGTCAGAGAAAGGACACTGG---ACATGTGTGCGCCATGAAGGTGTGTCGCAA 508

QY 3343 AGCGGAGCTTCTACGCGGATCAGGTGGCAGCAGCTGAAGGCCGAGAG---GGATATCCT 3399

Db 509 AGCGGACATGCTGGAAGAGGAGAGGTGGCAGAGTACGCGCGGAGGTCTGATGTCTCT 568

QY 3400 CGCGAAGCCGACAAATCACTGGTGGTGAAGTCTACTACAGTTCACGAGCAAGAGGATAA 3459

Db 569 GGTGAGCGCGATCATCAGTGGTGGTGAAGATGTACTACAGTTCCAGGATCCCGTCAA 628

QY 3460 TCTGTACTTTGTGATGACTACATACAGGTTGGTGTGATGTCGTGCTCATCAAACT 3519

Db 629 TTTATATTGATAATGAGTCTTTCGCTGGTGGTGTATGATGATGACGCTTTTAAATGAAGAA 688

QY 3520 GGGCATTTTCGAGAGAACTGGCCAGATTTCTACATCGCGGAGTCACTCGCGCTGGA 3579

Db 689 GGACAGCCTATCCGAGGAGGCGACACAGTTCTATATGAGTGAAGAGCGGATTTGGCGATCGA 748

QY 3580 CAGGTTTCACAAATGGCTTCATTCACAGAGACATCAAGCCCTGACAACTACTCATCGA 3639

Db 749 TTCTATTCACAACTCGTTTCATACACAGGATATCAAGCCCGATACATTCGTGCTGGA 808

RESULT 9

US-09-338-132-1

; Sequence 1, Application US/09338132

; Patent No. 6040164

; GENERAL INFORMATION:

; APPLICANT: Hemmings, Brian A.

; APPLICANT: Millward, Thomas A.

; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases

; FILE REFERENCE: 4-20265/A/PCT

; CURRENT APPLICATION NUMBER: US/09/338,132

; CURRENT FILING DATE: 1999-06-22

; EARLIER APPLICATION NUMBER: 08/860,150

; EARLIER FILING DATE: 1997-06-19

; EARLIER APPLICATION NUMBER: PCT/EP95/05052

; EARLIER FILING DATE: 1995-12-20

; EARLIER APPLICATION NUMBER: 94810746.1

; EARLIER FILING DATE: 1994-12-22

; NUMBER OF SEQ ID NOS: 18

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1

; LENGTH: 2101

; TYPE: DNA

; ORGANISM: Drosophila melanogaster

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (132)..(1499)

US-09-338-132-1

Query Match 4.68; Score 261.2; DB 3; Length 2101;

Best Local Similarity 55.08; Pred. No. 1.1e-52;

Matches 581; Conservative 3; Mismatches 451; Indels 21; Gaps 3;

QY 3043 CAAGTCTTCATGAGCAGACATAGAGACGTCATCAAGTCGTATGCCAGCGCAGTA 3102

Db 212 CAAGGTGACGTTGGAGAACTACTACAGCAACCTGGTGACGAGTATGGCGAGGAAAGCA 271

QY 3103 TCGCAAGAATCAGCTGGAGAGAGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAAT 3162

Db 272 GCGCTCGCAAGAGCTGGAGGCTCAGCTGAAGGAGGAGAGCTTGTTCGGAGGCGGAGCGCA 331

Fri Jan 17 11:17:10 2003

```

Db 1585 TACTCACCGTCGGTACGACACTACATCGCCCGGAGATCTTCTCTACAGGGCTAC 1644
Qy 3899 ACCAGCTGTGCGACACTACTGGAGCGTGGCGTCATCTCTTAYCAGATGCTGTGGTGGTAC 3958
Db 1645 GGTACAGGATGCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1704
Qy 3959 CCGCCCTTTCTGCGCAACAGTGGCTGGAAACGCAACAAAGGTCAATCAACTGGGAGAAA 4018
Db 1705 CCGCGGTCTGTCTGGAGACCCCGGAGGAACTACAGGAAGATCATGAATTCGAGCAG 1764
Qy 4019 ACSCTGCATATTCGCGCGCAGGCGGAGTATCCCGGAGGCTACGACATGATGATAGGAGG 4078
Db 1765 ACGTGTGTCTCCAGACGACATCCACATCTCATACGAGCAGAGGACCTCATCCCGCG 1824
Qy 4079 CTCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4135
Db 1825 CTGCTCTCGGACGCGGAGGAGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1884
Qy 4136 GACTTCTTCAAGGCGATCGACTTTGGCGACATCGGAGCAGAAAGGCGCTACATACCG 4195
Db 1885 CCGTCTTCCGCGGTGGACTGGGAGACCATCGCCAGGTGGCGGCTCCCTACATCCCG 1944
Qy 4196 GAATCAAGCACCCRAGGACACATCAACTTTGATCCCGTGGATCGGAGAAAGCTGC 4253
Db 1945 AAGCTGTCCAGCGTCACCGACACGCGCTTCTTCCCAACACAGAGCTGGAGACGTGC 2002

RESULT 11
US-860-150-6
; Sequence 6, Application US/08860150B
; Patent No. 5981205
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT FILING DATE: 1997-06-19
; EARLIER FILING DATE: 1997-06-19
; EARLIER FILING DATE: 1995-12-20
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (596)..(1990)
US-860-150-6

Query Match 4.1%; Score 232.6; DB 2; Length 3018;
Best Local Similarity 53.1%; Pred. No. 9.6e-46;
Matches 599; Conservative 2; Mismatches 496; Indels 30; Gaps 4;

Qy 3108 AGAATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATCGAGA 3167
Db 729 AAAAGAAGTTAGAAAAGGTGTAAGAGAGAGAGGCGCTAAAGATGAGGAGAAACGACTCC 788
Qy 3168 TGAGGAAATGCTGAACCAAAAGAGGACCACTACATTCGATTGAAGCGCGGCAAGATGG 3227
Db 789 GGAGATCAGCACATGCTCGGAGGAAACAGAGTTTCTGCTTTGAAGAGAAACAAAGACTTG 848
Qy 3228 ACAAGACATGTTGCTCAAACTAGACCCCATTTGGAGTGGTGCATTTGGCGAGGTAAACG 3287
Db 849 GATTGGAAGATTTGAGTCTCTTAAAGTAAATAGGACAGAGCATTTGGTGGAGTACGGC 908
Qy 3288 TGGTGACAAAATTCGATACCTCGAACCATTTGTATGCGATGAAACCCCTCGGGAAGCGG 3347
Db 909 TTGTCAGAGAAAGATA---CGGACATGTGTATGCAATGAAATTAATCTCCGTAAGACAG 965

```

```

Qy 3348 ACCTTCTCAACGGAATCAGTGGCAGACAGTGAAGCCGAGAGGATATCTCTCGCGAAG 3407
Db 966 ATATGCTTTGAAAAGACAGAGGTGGCCACATTCGTGCGGAGCGTGCATCTAGTGAGG 1025
Qy 3408 CCGACAATAACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATTAATCTGTACT 3467
Db 1026 CAGACAGTTTCTGGTGGTGAATAATGTTCTATAGTATTTTCAGGATTAAGCTTAACCTCTACC 1085
Qy 3468 TTGTGATGAGTACATACCAAGTGGTGTATGCTGTCTCATCAAACTGGGCAATTT 3527
Db 1086 TAATCATCGAGTTCTGCTGGAGGGACATGATGACCTTGTGTGAAAGAACACACTC 1145
Qy 3528 TCGAGGAGAACTGGCCAGATTTACATCGCGGAGGTACCTGCGCGGTGACAGCGTTC 3587
Db 1146 TGACAGAGAGGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTC 1205
Qy 3588 ACAAAATGGCTTCAATTCACAGAGACATCAAGCCCTGACAAACATCTCATCTCATAGGAGC 3647
Db 1206 ACCAACTTGGATTCATCCAGAGACATCAAAACAGACACTTCTTTTGGACAGCAAGG 1265
Qy 3648 GACACATAAAGCTCACGACTTTGGCCCTGTGCACGGGATTCGGATGGACGACAACTCGA 3707
Db 1266 GCATGTGAACTTCTGACTTTTGTGCTTTGACAGGACTGAAAAAGACATAGGACAG 1325
Qy 3708 AGTACTTACAGAGAACGCGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAGGAAT 3767
Db 1326 AATTTTA-TAGGAATCTGAACACACAGCCCTCCCGCATGATTTTCACCTTCCAGAACATGAAT 1384
Qy 3768 ACTCCGAGAACGCGGAGCCACCGTCTCTGGAGAGGCGACGGATGCGCGCATCAACAAA 3827
Db 1385 TCCAAAGGAAAGACAGAAACCTTGGAAAAAATAGAGTCA----- 1425
Qy 3828 GAGTCTCTGGCCCACTCGCTGGTGGGCAACCCCGAATACATAGTCTCCGAGGTGCTGGAGA 3887
Db 1426 ---GCTAGCCTTCTCCACAGTAGGACACTCTCTGAGGTGATCTCTGAGGTGTTTATGCG 1481
Qy 3888 GRAGTGGTACAGCGAGCTGTGCGACTACTTGGAGCGTGGCGCTCATCTCTTATGAGATCG 3947
Db 1482 AGACCGGTTACAAACAGCTCTGTGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1541
Qy 3948 TGGTGGGTACGCGCCCTTTCTTGGCCAAACAGTCCCGTGGAAACGCAACAAAGGTATCA 4007
Db 1542 TCATCGGCTACCCACCTTTCTGTTCTGAGACCCCTCAAGAGACATATAAGAGGATGATGA 1601
Qy 4008 ACTGGGAGAAACSCATATTCGCGCGAGCGCGAGTTATCCCGGAGGCTTACGAGCT 4067
Db 1602 ACTGGAAGAAACCTTTGACATTTCTCCCAAGAGTCCCATCTCTGAGAAAGCAAGATC 1661
Qy 4068 TGATAAGGAGGCT---CTGTGCGCTGGCTGACAAAGCGGTGGGCAAGAGCGTGGAGAGG 4124
Db 1662 TAATTTTGAAGTCTGCTGCTGTAATGGGAACATAGATTGGAGCTCTCTGAGTTGAGGAAA 1721
Qy 4125 TCAAGAGCCAGGACTTCTTCAAGGGCATCGACTTTGCGGACATGCGGAAACAGCAAGCGC 4184
Db 1722 TAAAGGTAACTCTTTTTTTTGAAGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1781
Qy 4185 CCTACATACCGGAATCAAGCACCCACGAGGACATCAACTTTGAT 4231
Db 1782 CAATATCTATTGAATCAAAAGCATTTGATGATACCTTCAAACTTCGAT 1828

```

```

; RESULT 12
US-09-338-132-6
; Sequence 6, Application US/09338132
; Patent No. 6040164
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 08/860,150

```

```

; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: PCT/EP95/05052
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: 94810746.1
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (596)..(1990)
US-09-338-132-6

Query Match
Best Local Similarity 4.18; Score 232.6; DB: 3; Length 3018;
Matches 599; Conservative 2; Mismatches 496; Indels 30; Gaps 4;

Qy 3108 AGAATCAGCTGAGAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCAAATCGAGA 3167
Db 729 AAAAGAACTTGAAGAGTGTGGAAGAGAGGCTTAAAGATGAGGAGAACACACTCC 788

Qy 3168 TGAGGAAATGCTGAACCAAGAGAGACACTACATTCGATTGAAGCGCGCCAAAGATGG 3227
Db 789 GGAGATCAGACATCTCGGAAGGAAACACAGATTTCTTCTTGAAGAGAAACAGACTTG 848

Qy 3228 ACAAGAGCATGCTCAAACTGAAGCCATTTGGAGTGGTGGTCAATTTGGCGAGGTAAAGC 3287
Db 849 GATTGGAAGATTTTGGTCTTAAAGTAATAGGAGAGGAGCAATTTGGTGAAGTACGGC 908

Qy 3288 TGGTGAGCAAAATCACTCGAACCACTTTGTATGCGATGAAACCTCGCGAAAGCGG 3347
Db 909 TTGTTCAAGAAAGATA---CGGACATGTGTATGCAATGAAATATCTCCGTAAGACAG 965

Qy 3348 ACGTTCTCAAGCGGAATCAGGTGGCGACACGTGAAGCGCGAGAGGATATCTTCGCGGAAG 3407
Db 966 ATATGCTTGAAGAGAGCAGGTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGG 1025

Qy 3408 CCGACAATTAAGTGGTGTGAAGTTGTACTACAGCTTCCAGACAGAGATATCTGTACT 3467
Db 1026 CAGACAGTTTGTGGGTGTGAAATGTTCTATAGTTTTCAGGATAAGTAACCTCTACC 1085

Qy 3468 TTGTGATGAGCTACATACAGGTGGTGTATCTGTCTGCTCATCAAACTGGGCAATTT 3527
Db 1086 TAATCATGAGTCTCTGCTGGAGGGACATGATGACCTTTGTATGAAAAAGACATC 1145

Qy 3528 TCGAGGAGAACTGCGCCAGATTTACATTCGCGGAGGTGACCTGCGCGCGTGGACAGCGTTC 3587
Db 1146 TGACAGAGAGGAGACTCAGTTTATATAGCAAGAAAGTATTAGCCATAGACTCTATTC 1205

Qy 3588 ACAAATGGCTTCAATTCACAGAGACATCAAGCTGACGACATCTACTCATAGAGGAGC 3647
Db 1206 ACCACTTGGATTCATCCAGAGACATCAACACAGACACCTTTTGGACAGCAAGG 1265

Qy 3648 GACACATAAAGCTCACCAGCTTTGCGCTGTGACGGGATTCGGATGGACGCAACTCGA 3707
Db 1266 GCCATGTGAACCTTTCTGACTTTGCTTTTCACAGGACTGAAAAACACATAGGACAG 1325

Qy 3708 AGTACTACAGAGAACGGCAATCACTCGCCAGGAGTCTGATGGAGCCCTGGGAGGAAT 3767
Db 1326 AATTTTA-TAGGAATCTGAACACAGACGCTCCCGAGTATTTCACTTTCCAGAACATGAAT 1384

Qy 3768 ACTCGGAGAACGGACCGGACCCACCTGCTGGAGAGCGGAGTGGCGCATCAACAAA 3827
Db 1385 TCCAAAGGAAAGCAGAAACCTGAAAAAGAAATAGACGTCA----- 1425

Qy 3828 GAGTCTCGGCACTCGCTGTGGGACCCCGAACTACATAGTCCCGAGGTGCTGGAGA 3887
Db 1426 ---GCTAGCCTTCTCCACAGTAGGCACTCTGACTACATGCTCTCGTGGAGTGTTCATGC 1481

Qy 3888 GAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGCGCTCATCTTAYGAGATGC 3947

```

```

RESULT 13
US-08-878-989-11
; Sequence 11, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/878,989
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:

```

```

Db 1482 AGACCGGTTACAAAGCTCTGTGATTTGGGTGCTGATCATGTATGAGATGC 1541
Qy 3948 TGGTGGGTACCGCGCTTTCTGGCCACAGTCCCTGGAACGCAAAAGGTATCA 4007
Db 1542 TCATCGGTACCCACCTTTCTGTTCTGAGACCCCTCAAGAGACATATAAGAGGTGATGA 1601
Qy 4008 ACTGGGAGAAAACCTGCATATTCGCGCGAGCGAGTTATCCCGCGAGGCTACGGACT 4067
Db 1602 ACTGGAAGAAACTTTGACTTTTCTCCAGAGTTCCCATCTCTGAGAAAGCCAGGATC 1661
Qy 4068 TGATAAGAGGCT---CTGTGCGTCTGACAAAGCGCTGGGCAAGACGCTGGACGAGG 4124
Db 1662 TAATTTTGGAGTTCGTGTTGAATCGGAACATAGAAATTTGGAGCTCCTCGAGTTGAGGAAA 1721
Qy 4125 TCAAGAGCCAGACTTCTTCAAGGCGATCGACTTTCCGAGACATGCGGAGACAGAAAGCGC 4184
Db 1722 TAAAAAGTAACCTCTTTTGAAGCGCTTGACTGGGAACATATCAGAGAGACCTGCTG 1781
Qy 4185 CTACATACCGGAATCAAGCACCCAGGACACATCCAACTTTGAT 4231
Db 1782 CAATATCTATTGAATCAAGACCATTTGATGATACCTCAACTTCGAT 1828

```

Qy	3043	CAAGTTCTTCATGGAGCAGCATAGAGAAAGTGATCAAGTCGTATCGCCAGCGCAGCTA	3100
Db	121	CAAGCTCACATGGGAATTTTTATAGCAACTAATTTTACAGCATGAAGAGAGAGAAAC	180
Qy	3103	TCCCAGAATCAGCTGGAGAAAGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAAT	3162
Db	181	CAGCGAGAAGAAATAGAAAGTGGCCATGGAAGAAGAGGATTACAGCATGAAGAGAAAA	240
Qy	3163	CGAGATGAGGAAATGCTGAACCAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAA	3222
Db	241	GTTCAGTCGATCACACACCGCTCGCAAGAAACAGAGTTCATACGGCTCAAAAGGACCAG	300
Qy	3223	GATGGACAAGACATGTTCGTCAAACTGAAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGT	3282
Db	301	ACTTGGCTTGGATGACTTTGAGTCTCTCAAAGTTATAGGAAGAGGAGCTTTTGGAGAGGT	360
Qy	3283	AAGCTGGTTCACCAAAATCGATACCTTCGAAACCATTTGTATCGGATGAAAACCCCTGGGAA	3342
Db	361	GGGTTGGTCCACAAAAGATACAGG---CCATATCTATGCAATGAAGATATTGAGAAA	417
Qy	3343	AGCGGACGTTCTCAAGCGGAATCAGGTGGCAGACGTGAAGCCGAGAGGATATCTCGC	3402
Db	418	GTCTGATGCTTGAANAAGACGAGTGGCCCATATCCGACGACAAAGATATTTGGT	477
Qy	3403	GGAGCCGACAATACTGGGTGGTGAAGTTGTACTACAGTTTCAGGACCAAGGATAACT	3462
Db	478	AGAAGCAGATGTGCTGGGTGGTGAAGATGTTTTACAGTTTTTCAGGATTAAGAGGAATCT	537
Qy	3463	GTACTTTGTGATGGACTACATACAGAGTGGTGATCTGATTCGCTGCTCATCAACTGGG	3522
Db	538	TTATCTAATCATGGAATTTCCCTTGGAGGTGACATGACATTCGTCAATGAAGAAAA	597
Qy	3523	CATTTTCGAGGAGGAACGTGCCAGATTACATCCCGAGGTCACTTCGCCCTGGGACAG	3582
Db	598	CACCTTGACAGAAGAGGAAACACAGTTCTACATTTTCAGAGACTGTTCTGCGCAATACATGC	657
Qy	3583	CGTTTACAAAATGGCTTCATTTACAGAGACATCAAGCCTGACAACTACTCATCGATAG	3642
Db	658	GATCCACCAAGTTGGGTTTCATCTCCATCGGGATATTAAOCCAGACAACCTTTATTTGGATGC	717
Qy	3643	GGACGGACACATAAAGCTCACCGACTTTGGCTCTGTCACGGGATTCGATGGAGCGACAA	3702
Db	718	CAAGGTCATGTAAATATTCTGATTTTGGTTTCAATGTACGGGATTAAGAAAGCTCACAG	777
Qy	3703	CTTGAAGTACTACCAGGAGAACGGCAATCACTCCGCCAGGATCGATGAGCCCTGGGA	3762
Db	778	GACTGAAATTTTATAGAAATCTCACACACAACCCACCAAGTGACTTCTCTATTCAGAAACAT	837
Qy	3763	GCATACTCCGAGACGACCGAAGCCCCACCGTGCTGGAGAGCGGAGTGGCGCATCA	3822
Db	838	GAACTCAAGAGGAAGACAGAAACTTGGAGAAGAACAGGAGACAA-----883	
Qy	3823	CCAAAGAGTCTTGGCCCACTCGCTGGTGGGACACCCCGAACTACATAGCTCCCGAGGTGT	3882
Db	884	-----CTGGCATATCCACAGTTGGGACACAGATTACATTTGCTCCAGAGATT	933
Qy	3883	GGAGAGAGTGGTACCGCAGCTTCGCGACTACTGGAGGTGGGCGTCATCTCTTAYGA	3942
Db	934	CATGCAGACTGGTTACACAAAATTTGTGTGACTGGTGTCTTTGGGAGTGATTATGTATGA	993
Qy	3943	GATGCTGTGGTCAAGCGCCCTTTCTTGGCCACAGTCCGCTGGAACCGACAAAGCT	4002
Db	994	ATGCTAATAGGATATCCACCTTTCTGCTGTGAACACCTCAAGAAACATACAGAAAAGT	1053
Qy	4003	CATCAACTGGGAGAAAACSTGTCATATTTCGCCGCGAGCGCGAGTTATCCGGGAGGCTAC	4062

Matches 609; Conservative 2; I

Gaps : 3;

QY 3043 CAAGTCTTTCATGGAGCAGCATAGAGAACTGATCAAGTCTGATCGCCAGCCAGCA 3102
 DB 121 CAAGCTCATTTGGAGAAATTTTATAGCAACTAATTTTACAGCATGAAGAGAGAGAA 180
 QY 3103 TCGAAGATCAGCTGGAGAGAGATGCAAAAGTGGGACTGCCGATCAGACCAAT 3162
 DB 181 CAGGCAAGAAATTTAGAGTGGCCATGGAAGAAGGATTTAGCAGATGAAGAGAAA 240
 QY 3163 CGAGATGAGAAATGCTGAACCAAAAGGAGAGCAACTACATTCGATTGAAGCGGCCAA 3222
 DB 241 GTTAGCTGCATCAACACAGCTCGCAAGAAACACAGTCTTACGGCTCAAAAGACCA 300
 QY 3223 GATGACAAGAGCATGTTGCTGAACCTGAAGCCCATTTGGAGTGGTGGTATTTGGGAGGT 3282
 DB 301 ACTTGGCTGGATGACTTTGAGTCTCTGAAAGTTATAGGAAGAGGAGCTTTTGGAGAGT 360
 QY 3283 RACGCTGGTGACCAAAATCGATACCTCGAACCATTTGATGGGATGAACACCTCGGAA 3342
 DB 361 GCGGTGGTGCCACAAAAGATACAGG---CCATATCTATGCAATGAAGATATTGAGAAA 417
 QY 3343 AGCGACGCTTCTCAAGCGGAATCAGGTGGCAGACGTTGAAGCGGAGAGGATATCTCGC 3402
 DB 418 GTCGATATGTTGAAAGAGCAGCTGGCCATATCGAGCAGAGAAAGATATTTGGT 477
 QY 3403 GGAAGCCGACATAACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATAATCT 3462
 DB 478 AGAAGCAGATGGTGGTGGTGAAGATGTTTACAGTTTTCAGGATAGAGGAATCT 537
 QY 3463 GTACTTTGTGATGATACATACAGGTGGTGGTATCTGATGCTGCTCATCAAACTGGG 3522
 DB 538 TTATCTAATCATGAAATTTCTCCCTGGAGGTGACATGATGACATTTGCTAATGAAGAAA 597
 QY 3523 CATTTTTCAGGAGGAGTGGCCAGATTTCTACATCCCGAGGTCACTCGCGCGTGGACAG 3582
 DB 598 CACCTTGACAGAGAGGAAACACAGCTTCTACATTTACAGAGACTGTTCTGGCAATAGATGC 657
 QY 3583 GATTCAAAAATGGCTTTCATTCACAGAGACATCAAGCTGCAACATCTCATCTGATAG 3642
 DB 658 GATCCACAGTGGTGGTTCATCCATCGGATATTAAGCCAGACACCTTTTATTGGATGC 717
 QY 3643 GGAGGACACATAAGCTTCACCGACTTTGGCTGTGACGGGATTCGATGGACGACAA 3702
 DB 718 CAAGGTCATGTAATATTTGATTTTGGTTCATGTACGGGATTAAGAAAGCTCACAG 777
 QY 3703 CTCGAAGTACTACAGGAGAACGGCAATCACTCGCGCGAGGACTCGATGGAGCCCTGGGA 3762
 DB 778 GACTGAATTTATAGAAATCTCACACACACCCCAAGTACTCTCATTTCAAGACAT 837
 QY 3763 GGAATACCTCGAGAACGGACCGAACCCCGTGTGAGAGGGCGACGGATCGCGATCA 3822
 DB 838 GAACCTAAAGAGGAAAGCAGAACTTGAAGAAGAAACAGGAGACAA----- 883
 QY 3823 CCAAGAGTCTCGCCACTCGCTGGTGGGACCCCGCACTACATAGTCTCCGAGGTGCT 3882
 DB 884 -----CTGGCATATTCACAGTTGGGACACACGATTTACATTTGCCAAGAGTATT 933
 QY 3883 GGAGAGAGTGGGTACACGAGCTGTGCGACTACTGTGAGCGGTGGCGTCTATCTTATGA 3942
 DB 934 CATGAGACTGGTTACACAAATTTGTGTGACTGGTGGTCTTTGGAGTGATTTATGATGA 993
 QY 3943 GATGCTGGGTGAGCGCCCTTTCTGGCCCAACAGTCCGCTGGAAACGCAACAAAGGT 4002
 DB 994 AATGCTAATAGGATATCCACCTTTCTGCTCTGAAACACCTCAAGAAACATACAGAAAAGT 1053
 QY 4003 CATCACTGGGAGAAACCTTCATATTCGCGCGGAGCGGAGTTATCCGGGAGGCTAC 4062
 DB 1054 GATGAACCTGGAAGAAACCTCTGGTATTTCTCCAGAGGTACCTATATCTGAGAAAGCCAA 1113
 QY 4063 GGACTTGAAGAGGCTCTGTGCGCTGACAAAGCGGTGGGCAAGA---GGTGGGA 4119
 DB 1114 GGACTTAATCTCAGATTTTGTATTGATTTCTGAAACAGAAATTTGGAATAGTGGAGTGA 1173
 QY 4120 CGAGTCAAGAGCCACGACTTCTTCAAGGGGATCGACTTTTGGGACATCGCGAAGCAGAA 4179

DB 1174 AGAAATAAAGTGCATCCCTTTTGAAGTGTGCTGAGTGGGACACATAAGGAAAGGCC 1233
 QY 4180 AGCGCCCTTACATACCGGAAATCAAGCACCCACGAGGACACATCCAACTTTGATCCCGT 4236
 DB 1234 AGCAATCCCTATAGAAATCAAAAGCATTTGATGATATCTTCAAAATTTTGTGACTT 1290

RESULT 15
 US-09-328-111-66
 ; Sequence 66, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astie, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; TITLE OF INVENTION: PRODUCTS
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; CURRENT FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; EARLIER FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 66
 ; LENGTH: 678
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(678)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-328-111-66

Query Match 2.9%; Score 164.6; DB 4; Length 678;
 Best Local Similarity 63.2%; Pred. No. 7.7e-30;
 Matches 289; Conservative 0; Mismatches 161; Indels 5; Gaps 2;

QY 3078 TCAAGTCTGATCGCCAGCGCAGCTATCGCAAGAATCAGCTGGAGAGGAGATGCACAAG 3137
 DB 3 TCAATCTCATCAGCAGCGCTCTACATCGTAAACAAATAGAGAATGAATGTCGGG 62
 QY 3138 TGGGACTCCCGATCAGACCCCAATCGAGATGAGGAAATGCTGAACCAAAAGGAGAGA 3197
 DB 63 TTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAACTA 122
 QY 3198 ACTACATTCGATTGAAGCGCGCAAGATGGCAACAGAGATGTTCTGTCAAACTGAAGCCCA 3257
 DB 123 ATTACATCCGCTTAAAAGGGCTAAATGGCAAGTCTATGTTGTGAAGATAAAGACAC 182
 QY 3258 TTGGAGTGGGTGCATTTTGGCGAGGTAAAGCTGGTGAGCAAAATCGATACCTCGAACCAATT 3317
 DB 183 TAGGAATAGGAGCATTTTGGTGAAGTCTGCTAGCAAGAAAGTAGATA--CTAAGGCTT 239
 QY 3318 TGTATGCGATGAACCCCTCGGGAAGCGGAGCTTCTCAAGCGGAATCAGGTGGCACAG 3377
 DB 240 TGTATGCAACAAAACCTCTCGAAAGAAAGATGTTCTTCTTGGAAATCAAGTCGCTCATG 299
 QY 3378 TCAAGCGCGAGGAGATATCTCGCGGAAGCGGACCAATTAACCTGGGTGGTGAAGTTGTACT 3437
 DB 300 TTAAGGCTGAGAGAGATATCTTGGCTGAAGCTGACATGAATGAATGGGTAGTCTGCTATATT 359
 QY 3438 ACAGCTTCCAGGACAAAGGATAATCTGTACTTGTGTGAGACTACATACCAAGCTGGTGATC 3497

Db 360 ATTCATTCCAAGATAAGGCCATTATCCTT-GTAATGGCTACATTCCTNGGGTGATA 417
QY 3498 TGAATGCTGCTGCTCATCAAACTGGGCATTTT 3528
Db 418 TGAAGAGCCCAATTAATANAATGGGCATCTT 448

Search completed: January 16, 2003, 21:59:35
Job time : 245.379 secs

GenCore version 5.1.3.
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:42 ; Search time 158.731 Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1' atctgacacgagcgagca.....gccttgccaattgattac 5720

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	487.8	8.5	2043	9	US-09-836-392-2
2	321.4	5.6	676	9	US-09-764-868-214
3	273.6	4.8	638	10	US-09-879-536-26
4	232.6	4.1	3583	9	US-09-974-238-152
5	200.6	3.5	1452	9	US-09-938-842A-2402
6	195.8	3.4	1689	9	US-09-938-842A-1861
7	188.6	3.3	1818	10	US-09-771-161A-89
8	164.6	2.9	678	10	US-09-879-536-66
9	161.2	2.8	734	9	US-09-764-868-196
10	144.6	2.5	1398	9	US-09-938-842A-633
11	138.6	2.4	1416	9	US-09-938-842A-2503
12	134.6	2.4	3061	10	US-09-880-107-2146
13	134.6	2.4	3407	10	US-09-971-845-1
14	133	2.3	362	9	US-09-796-692-7789
15	127	2.2	1635	10	US-09-880-107-2340
16	121.8	2.1	1244	10	US-09-771-161A-38
17	121.8	2.1	1393	10	US-09-771-161A-37
18	121	2.1	1735	9	US-09-764-868-58
19	120.6	2.1	2549	10	US-09-880-107-3691

20	117	2.0	2556	10	US-09-817-310-1	Sequence 1, Appli
21	116.6	2.0	1257	10	US-09-799-875-15	Sequence 15, Appl
22	116.6	2.0	1826	10	US-09-799-875-13	Sequence 13, Appl
23	116.6	2.0	1972	9	US-10-098-841-166	Sequence 166, App
24	115.4	2.0	2146	10	US-09-954-456-527	Sequence 527, App
25	115.4	2.0	2146	10	US-09-842-307-1	Sequence 1, Appli
26	114.4	2.0	1479	10	US-09-771-161A-46	Sequence 46, Appl
27	113.4	2.0	2637	10	US-09-799-875-3	Sequence 3, Appli
28	113.4	2.0	3003	10	US-09-799-875-1	Sequence 1, Appli
29	113	2.0	277	10	US-09-294-093B-2043	Sequence 2043, Ap
30	110.8	1.9	568	9	US-09-796-692-8122	Sequence 8122, Ap
31	110.8	1.9	6409	10	US-09-864-864-293	Sequence 293, App
32	110	1.9	2610	10	US-09-970-000-3	Sequence 3, Appli
33	104.8	1.8	1812	10	US-09-971-118-1	Sequence 1, Appli
34	104.6	1.8	1461	10	US-09-771-161A-63	Sequence 63, Appl
35	104.6	1.8	1515	10	US-09-804-471A-1	Sequence 1, Appli
36	104.6	1.8	5877	12	US-10-028-946-3	Sequence 3, Appli
37	104.6	1.8	6165	12	US-10-028-946-1	Sequence 1, Appli
38	104	1.8	3244	9	US-10-174-590-571	Sequence 571, App
39	104	1.8	3244	9	US-10-176-758-571	Sequence 571, App
40	104	1.8	3244	12	US-10-052-586-571	Sequence 571, App
41	103.8	1.8	6303	9	US-09-974-238-168	Sequence 168, App
42	102.8	1.8	512	10	US-09-867-701-6230	Sequence 6230, Ap
43	102.2	1.8	2519	9	US-10-071-766-13	Sequence 13, Appl
44	102.2	1.8	2557	10	US-09-972-694-1	Sequence 1, Appli
45	101.4	1.8	678	10	US-09-841-683-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match	8.5%	Score 487.8;	DB 9;	Length 2043;
Best Local Similarity	65.2%	Pred. No. 1.3e-97;		
Matches 819;	Conservative 4;	Mismatches 401;	Indels 33;	Gaps 6;
QY	3130	GCACAAAGTGGGACATCGCGGATCAGACACCAATCGAGATGAGGAAATGCTGCAACAAA	3189	
DB	211	CGGTAATGCTGCGACTCTGTGAAGCTTGACGAGGAGGATCGGGAAGATCCTCTACAGAA	270	
QY	3190	GGAGAGCACTACTTCATTCGATTCGAGCGCCCAAGATGACACAGCATGCTTCGTCAAACT	3249	
DB	271	AGAGTCTTAATACACAGGTTAAAGAGGCCCAAGATGACAACTGTTGTTCAAGT	330	
QY	3250	GAAGCCCATGAGGTGGTGCATTTGGCGAGGTAAAGCTGGTGGAGCAAAATCGATACCTC	3309	
DB	331	CAAAACCCCTGGGATCGGTCCTTTGGAGAAAGTGTGCTTGTAGGTGGA---CAC	387	

QY	3310	GAACCA	TTTGTATGCGATGAAACCCCTGCGGAAGCGGAGCTTCTCAACGCGGAATCAGGT	3369
DB	3388	TCACGCCCTCTAGCCCATGAAGACCCCTAAGAAAAAGGATGTCTTGAACCGGAATCAGGT	447	
QY	3370	GGCACACGCTGAAGCGCCGAGAGGATATCTCTCGCGGAAGCGGACAATACTGGGTGTGAA	3429	
DB	448	GGCCACGCTCAAGCGCCGAGAGGACATCTCTGGCGAGGACAGCAATGAGTGGGTGTGCA	507	
QY	3430	GTTGTACTACAGCTTCAGACACAGGATATCTGTACTTTGTGATGGACTACATCCCTGG	3489	
DB	508	ACTCTACTACTCTCTTCAAGACAAGACAGCCTGTACTTTGTGATGGACTACATCCCTGG	567	
QY	3490	TGGTGATCTGATGTGCGTCTCATCAAACTGGGCATTTTCGAGAGAGAACTGCCAGATT	3549	
DB	568	TGGGCACATGATGAGCCTGTGCTGATCCGGATGGAGGTCTTCCCTGAGCACCTGGCCCGGTT	627	
QY	3550	CTACATGCCCGAGGTCACTTCGCCGTGGACAGCGTTTCAAAAATGGCTTTCATTCACAG	3609	
DB	628	CTACATCGCAGAGCTGACTTTTGGCCATGTGAGAGTGTCCACAAATGGCTTTCATTCACCG	687	
QY	3610	AGACATCAAGCCTGACAACATACTCATCGATAGGAGCGGACACATAAAGCTCACCGACTT	3669	
DB	688	AGACATCAAGCCTGTATACATTTTCATAGATCTGAGTGTACATTAACATTCACAGATT	747	
QY	3670	TGGCCTGTGACGCGGATTCGGANTGACGCCACAACCTCGAAGTACTACCCAGGA-----GAA	3723	
DB	748	CGGCCCTGCACTGGGTTCAGGTGGACTCACAAATTCCTCAAAATATTACCAAGAAGGAGCCA	807	
QY	3724	CGGCATCACTCGCCAGGACTCGATGGAGCCCTGGGAGGAATACTCCGA-----GAA	3777	
DB	808	TGTCAGACAGGACGATGGACCCAGCGACCTCTGGGATGATGTCTAACTGTCTGGTG	867	
QY	3778	CGSACCGAAGCCACCGTCTCGGAGGCGGACGGATGCGGATCACCAAGAGTCTCTGGC	3837	
DB	868	TGGGGACAGGCTGAAGACCCTTAGCAGAGGCGCGGAAGCAGCAGAGTGCCTGGC	927	
QY	3838	CCACTCGCTGGTGGGCACCCGAACTACATAGTCTCCGAGGTGCTGGAGAGAGTGGGTA	3897	
DB	928	ACATTTCACTGGTGGGACTCCAACTACATCGCACCCGAGGTCTCTCCGCAAGAGGTA	987	
QY	3898	CACGACGCTGGCAGCTACTGGAGCTGGCGGTCTATCTTAYGAGATGCTCGTGGGTCA	3957	
DB	988	CACCTCAACTGTGTACTGGTGGAGTGTGGAGTGATTTCTTCGAGATGCTGGTGGGCA	1047	
QY	3958	GCGCCCTTTCTGGCCAAACAGTCCGCTGGAAGCGCAACAAAAGTTCATCACTGGGAGAA	4017	
DB	1048	GCGCCCTTTTGGCACCTACTCCCACAGAAACCCAGCTGAAGTGTATCACTGGGAGAA	1107	
QY	4018	AACSGTCATATTCCCGCGCAGGCGGATATCCCGGAGGCTACCGACTTGATAGGAG	4077	
DB	1108	CACGCTCCACATTCACGCCAGGTGAGCTTGAGCCCTTGAGCGCAGGACCTCATCACAA	1167	
QY	4078	GCTCTGTGCGTCCGCTCACAAAGGCGTGGCAAGAGCGTGGCAGCGG-----TCAAGAGCCA	4134	
DB	1168	GCTGTGCTGCTCCGACACCAACCGCTTGGGCGGAATGGGCGGATGACCTGAAGGCCCA	1227	
QY	4135	CGACTTCTTCAAGGCGATCGACTTTGC-----GGACATCGGAAGCAGNAACGCGCCCTACAT	4191	
DB	1228	CCCCCTTCTCAGCGCCATTTGACTTCTCCAGTGTACATCCGAAGCAGCGCCCTACGT	1287	
QY	4192	ACCGGAATCAAGCACCCACGACACATCCAACTTTTGTATCCGTGATCCGAGAGCT	4251	
DB	1288	TCCCCACCATAGCACCCCATTTGGACCTCGAATTTGCACCCGTAGATGAAGAAGGCC	1347	
QY	4252	CGGCTCGAATGACTCCACCATGAGCAGCGCGGATGATGTC-----GACCAGAA	4299	
DB	1348	TTGGAACGATGCCAGCGAAGGTAGCACCAAGCGCTGGGACACACTACCTTCGCCCAATA	1407	
QY	4300	TGACCGCAYTTCACGCGCTTTTTCGAATTTTACCTTCGTCGCTTCTTCGACGACAA	4356	
DB	1408	CAAGCATCTCAGCAGCATTTTTCGAATTCACCTTCGAGGTTCTTTGATGACAA	1464	

```

RESULT 2
US-09-764-868-214
; Sequence 214, Application US/09764868
; Patent No. US20020168711AI
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (628)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-214

Query Match          5.6%; Score 321.4; DB 9; Length 676;
Best Local Similarity 69.7%; Pred. No. 2.8e-61;
Matches 474; Conservative 2; Mismatches 199; Indels 5; Gaps

QY 3041 TTCAAGTTCTTCATGAGCAGACATAGAGAACGTGATCAAGTCGTATCGTCCGACGCGACG 3100
DB 1 TTTAAGTTCTTCATGAGCAGCAGCTGGAGATGTCTCAAACCTCACACGAAAGTTT 60
QY 3101 TATCGCAAGNATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAA 3160
DB 61 AACCGGAGGCTGCAGCTGGAGCAAGNATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120
QY 3161 ATCGAGATGAGGAAATGCTGAACCAAAAGGAGACGAACTACATTCGATTTGAAGCGGCC 3220
DB 121 GACGAGATCGGAAGATCCTTACCAGAAGAGTCTAATTACACAGGTTAAAGAGGCC 180
QY 3221 AAGATGGCAAGAGCATGTTTCGTCAAACCTGAAGCCCATTTGGAGTGGGTGCTATTTGGCGAG 3280
DB 181 AAGATGGCAAGTCTATGTTGTTCAGATCAAAACCTCGGGATCGGTGCTCTTGGAGAA 240
QY 3281 GTAACGCTGGTGAGCAAAATTCGATACCTCGAACCTTTGTATCGCATGAAACCTCGCG 3340
DB 241 GTGTGCTTCTGTTGAAGTGGA--CACTCAGCCCTGTACGCCATGAAGACCTTAAGG 297
QY 3341 AAACGGCAAGTTCTCAGCCGATCAGTGGCAGACGTTGAAGCCCGAGAGGATATCCTC 3400
DB 298 AAAAAGGATGTCTTGAACCGGATCAGTGGCCACGTCAGTGGCCGAGAGGACATCCTG 357
QY 3401 CGGGAAGCCGACAATACTGGGTGGTGAAGTTGTACTACAGTTTCAGGACAGAGGATAAT 3460
DB 358 GCCGAGGACAGCAATGATGGTGGTGGCAAACTC-TACTACTCTCTCCAAGACAANGACAGC 416
QY 3461 CTGTACTTTGTGATGGACTACATACACAGGTGGTGTATGTATGTCGCTGTCTATCAAACTG 3520
DB 417 CTGTACTTTGTGATGGACTACATCCCTTGGTGGGACATGATGAGCCTGCTGTATCCGGATG 476
QY 3521 GGCATTTTCGAGGAGGAACCTGGCCAGATTCTACATCGCCGAGGTACCTCGCCGCTGGAC 3580
DB 477 GAGGTCTTCCCTGAGCAGCTTGGCCGGKTCCTACATCGCACA-CTGACTTTGGGCCATTGAG 535
QY 3581 AGCGTTTCAAAATGGGCTTCTATTCACAGACATCAAGCTGACACATCACTACTCATCGAT 3640
DB 536 AGTGTCACAGATGGGCTTCATCCCGCAGACATCAAGCCTGATACATTTTGTATAGAT 595
QY 3641 AGGAGCGGACATAAAGCTCACCGACTTTGGCTGTGACAGGATTCGATGACGCCAC 3700
DB 596 CTGGATGCTACATTAACATCACAGATTTCCGCTCTCACTGGGTTCAGTGGACTCAC 655
QY 3701 AACTCGAAGTACTACCAGGA 3720

```

Db 656 AATTTCACATATTACCAGAA 675

RESULT 3

US-09-879-536-26/c

; Sequence 26, Application US/09879536

; Patent No. US20020144298A1

GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

; APPLICANT: Bushnell, Steven E.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Catino, Theodore J.

; APPLICANT: Derti, Adrian

; APPLICANT: Ford, Donna M.

; APPLICANT: Lewis, Marcia E.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS

; FILE REFERENCE: CCD-257 (US)

; CURRENT APPLICATION NUMBER: US/09/879,536

; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: US 60/088,801

; PRIOR FILING DATE: 1998-06-10

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 638

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-879-536-26

Query Match 4.8%; Score 273.6; DB 10; Length 638;
Best Local Similarity 66.4%; Pred. No. 8.8e-51;
Matches 425; Conservative 0; Mismatches 209; Indels 6; Gaps 2;

QY 3074 GTGATCAAGTCGTATCGCCGCGGACGATATCCAGATACAGTCGAGGAGAGATGCAC 3133

Db 637 GTATCAAAATCTCATCAGCAGCGGTACATCGTAAACAAATAGAGAATGAAATGATG 578

QY 3134 AAATGGGACTCCCGATCAGACCAATCGAGATGAGGAAATGCTGAACCAAAAGAG 3193

Db 577 CGGTGGATATCTCAAGATGCCAGGATCAATGAGAGAGATGCTTGGCCAAAAGAA 518

QY 3194 AGCAACTACATTCGATGAAGCGGCCCAAGATGGACAGAGATGTCGTCAAACTGAAG 3253

Db 517 TCTAATTACATCCGCTCTAAAGGGCTAAATGGACAAGTCTATCTTTGTGAAGATAAG 458

QY 3254 CCCATTGGAGTGGTGCATTTGGCGAGGTAAAGTGGTGGAGCAAAATCGATACCTCGAC 3313

Db 457 ACATGAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAAAGTAGATA---CTAAG 401

QY 3314 CATTTGTATGCGATGAAACCCCTGCGGAAAGCGGACGTTCTCAAGCGGAATCAGGTGGCA 3373

Db 400 GCTTTGTATGCAACAAACTCTTCAATGAAGATGTTCTTTCGAAATCAAGTCGT 341

QY 3374 CACGTGAAGCGGAGGAGATATCTCGGGAGCGGACAAATCAATCGGTGGTGAAGTTG 3433

Db 340 CATGTAAAGGCTGAGAGATATCTGCTGAAGCTGACAAATGAATGGTGTCTGCTA 281

QY 3434 TACTACAGCTCCAGGACAGGATATCTGTACTTTGTGATGAGTACATACACAGTGGT 3493

Db 280 TATTATTCATCAAGATAAGGACAAATTTATCTTTGTAATGGACTACATTCCTGGGGT 221

QY 3494 GATCTGATGTCGCTGCTCATCAAACTGGGCATTTTCGAGGAGGAACTGGCCAGATTCTAC 3553

Db 220 GATATGATGAGCCTATTATTAATGATGGCATCTTCCAGAAAGTCTGCCAGCATTCATC 161

QY 3554 ATCGCCGAGGTACCTGCGCGCTGGACAGCG---TTCACAAAATGGGCTTCATTCACAGA 3610

Db 160 ATAGCAGAACTTACCTGTGCAGTTTGAAGGTGTTCCCTTAAAAATGGTTTTATTATCATA 101

QY 3611 GACATCAAGCCCTCACAACATACATCGATGAGGAGGACACATAAAGCTCACCACCTTT 3670

Db 100 GATATTAACCTGATAAATTTTGTATGATCGTGATGTCATATTAATTAATGACTGACTTT 41

QY 3671 GGCCTGTGTCACGGGATTCGATGGAGCGACAACTCGAAGT 3710

Db 40 GGCCTCTGCATGGCTTCAGATGGACACACAGGATTCAGT 1

RESULT 4

US-09-974-298-152

; Sequence 152, Application US/09974298

; Patent No. US20020156263A1

GENERAL INFORMATION:

; APPLICANT: Chen, Hui-Mei

; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER

; FILE REFERENCE: PA-0037 P

; CURRENT APPLICATION NUMBER: US/09/974,298

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: 60/238,331

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 152

; LENGTH: 3583

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20020156263A1 347975.11

US-09-974-298-152

Query Match 4.1%; Score 232.6; DB 9; Length 3583;
Best Local Similarity 53.1%; Pred. No. 2.2e-41;
Matches 599; Conservative 2; Mismatches 496; Indels 30; Gaps 4;

QY 3108 AGAATCAGCTGAGAGGAGATGCACAAAGTGGGACTGCCCATCAGACCCCAAAATCGAGA 3167

Db 417 AAAGAAGTTAGAAAGGTGATGGAAGAAGAGGCTTAAAGATGAGGAGAAACGACTCC 476

QY 3168 TCAGCAAAATGCTGAACCAAAAGAGAGACACTACATTCGATTGAAGCGCGCAAGATGG 3227

Db 477 GGAGATCAGCATGCTCGGAAGGAAACAGAGTTTCTGTTTGAAGAGAAACAGACTTG 536

QY 3228 ACAAGAGCATGTTGTCRAAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAACGC 3287

Db 537 GATTGGAAGATTTTTCAGTCCGTAAAAAGTAATAGGCAGAGGAGCAATTTGTTGAGGTACGGC 596

QY 3288 TGGTCAGCAAAATCAGTACCTCGAACCATTTGTATGCGATGAAACCCCTGCGGAAGCGG 3347

Db 597 TTGTTTCAGAGAGAGATA---CGGACATGTGTATGCAATGAAATACTCTCCGTAAGCAG 653

QY 3348 AGTTTCTCAAGCGGAATCAGGTGGCACACGTCGAAGCGCGAGAGGATATCTCTCGCGGAG 3407

Db 654 ATATCTTGAAGAAACAGCAGGTTGCGCACATTCGTCGAGGCGTCACATTTCTAGTGGAGG 713

QY 3408 CCGCAAAATACCTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATATCTGTACT 3467

Db 714 CAGACAGTTTGTGGTGGTGTGAAAATGTTCTATAGTATTTTCAGGATAGAGTAAACCTCTACC 773

QY 3468 TTGTGATGAGCTACATCAGGTGGTGTGATCTGCTGCTCATCAAACTGGGCATTT 3527

Db 774 TAATCATGAGTTCCTGCTGGAGGGGACATGATGACCTTGTGTATGAAAAGACACTC 833

QY 3528 TCGAGGAGAACTGCCAGATTTCTACATCGCGAGGTCACCTCGCCGCTGGACAGCGTTC 3587

Db 834 TCACAGAGAGGAGACTCAGTTTATATAGCAGAAACAGATATTACCCATAGACTCTATTC 893

QY 3588 ACAAAATGGGTTTCAATCAGAGAGACATCAAGCCCTGACAACATCTCATCGATAGGGAGC 3647

Db 894 ACCAACTTGGATTCATCCACAGAGACATCAAAACAGACAAACCTTCTTTTGGACAGCAAGG 953
 Qy 3648 GACACATAAAGCTCACCGACTTTGGCTGTGCGACGGGATTCGATGAGCGACCAACTCGA 3707
 Db 954 GCCATGTGAAGCTTTCGTACTTTGGTCTTTGGACAGGACTGAAAGAGACATAGAGACAG 1013
 Qy 3708 AGTACTACACAGGAGAGCGCAATCACTCGCGCCAGGAGCTGGATGGAGGCAAT 3767
 Db 1014 AATTTTA-TAGGAATCTGAACACAGCTCCCGAGTGATTTCACTTTCCAGAACATGAT 1072
 Qy 3768 ACTCCGAGAACGAGCAAGCCACCGCTGCTGGAGAGGCGCGGATCCGCCATCAACAAA 3827
 Db 1073 TCCAAAGGAAACAGCAAAACCTTGGAAAGAAATAGAGGTCA----- 1113
 Qy 3828 GATCTCGGCCACTCGCTGGTGGGACCCCGCAACTACATAGCTCCCGAGGTGCTGGAGA 3887
 Db 1114 ----GCTAGCTTCTCCACAGTAGGCACCTCTGACTACATTTGCTCTGAGGTGTTCAATGC 1169
 Qy 3888 GRAGTGGGTACAGCGAGCTGTGCGACTACTGGAGCGGTGGCGGTCTATCTCTTTAYAGATGC 3947
 Db 1170 AGACCGGGTACAAAGCTCTGTGATTTGGTGTGCTGGGTGATCATGTATGAGATGC 1229
 Qy 3948 TGTGGGTGAGCGCCCTTTCTGGCAACAGTCCGCTGGAAACGCAACAAAGGTCAATCA 4007
 Db 1230 TCATCGCTACCCACCTTCTGCTGTGAGACCCCTCAAGAGACATATAGAGGTGATGA 1289
 Qy 4008 ACTGGGAGAAACSTGCTATATTCGCGCCAGCGCGAGTTATCCCGAGGCTACGGACT 4067
 Db 1290 ACTGGAAAGAACTTTGACTTTTCTCCAGAAAGTTCCTATCTCTGAGAAAGCCAGGATC 1349
 Qy 4068 TGATAAGGAGGCT---CTGTGCGTGGCTGACAAAGCGCTGGCAAGCGCTGGCAAGCGTGCAGG 4124
 Db 1350 TAATTTTGGAGTCTGCTGTGATGGGACATAGAAATTTGGAGCTCCTGGAGTTGAGGAAA 1409
 Qy 4125 TCAAGGACCAACACTTCTTCAAGGCGATCGACTTTGGCGACATGCGGAAGCAGAAAGCGC 4184
 Db 1410 TAAAGTAAGTCTTTTGAAGCGTGTGACTGGGAACATATCAGAGAGAGACCTGCTG 1469
 Qy 4185 CCTACATACCCGGAATCAAGCACCCACGACACATCCAACTTTGAT 4231
 Db 1470 CAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTCGAT 1516

RESULT 5

US-09-938-842A-2402
 ; Sequence 2402, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938, 842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: 2001-08-24
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 2402
 ; LENGTH: 1452
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-2402

Query Match 3.5%; Score 200.6; DB 9; Length 1452;
 Best Local Similarity 49.6%; Pred No. 1.5e-34;
 Matches 640; Conservative 3; Mismatches 617; Indels 31; Gaps 4;

Qy 2984 AAGGAGAGGAGGAGGAGGAGCGCAAGGATTCGCGCATCAGCAGTACTCGCCGCAAGCCCTTC 3043
 Db 25 AACGCTAGCGGAGGAGGAGGTTCTGGATCGAGCTTGACCATGGAGAAAGTGGCGCAGCG 84
 Qy 3044 AAGTTCTTCATGGAGCAGCAGATAGAGAAAGTATCAAGTCTGATTCGCGCAGGACGCTAT 3103
 Db 85 AAGCAGTACATCGAGAAATCACTACAAAGCTCAGAACAGAAACATTTCAAGAGAGGAAAGAG 144
 Qy 3104 CCAAGAAATAGCTGGAGAGGAGATGCACAAAGTGGGAGCTGCCGATCAGACCCCAATC 3163
 Db 145 AGACGGTGGATCTTAGAAAGAAAGTTGGCTCTTCTGGAGTGCCAAAGAGGAGCAAAATC 204
 Qy 3164 GAGATGAGGAAATGCTGTAACCAAGAGAGAGCAATACATTCGATTGAAGCGCGCCCAAG 3223
 Db 205 AACATGATTAAGAGACCTGGAGAGAAAGAGACAGAGTTTATGAGGCTTAAAGAGAAACAAG 264
 Qy 3224 ATGGACAAGAGCAGTGTTCGTCACAACTGAAGCCCATTCGAGTGGGTGCTATTTGGGAGGTA 3283
 Db 265 ATTAGTGTGATGACTTTGAGCTTTGACTATCATTTGGAAGAGGTGCTTTTGGTGAGGTT 324
 Qy 3284 ACGTGTGTGAGCAAAATCGATACCTCGAACCATTTGTATGCGATGAAAAACCCCTCGCGAAA 3343
 Db 325 CGCTTATCTCGGAGAGAAAGTCTGGAA---ATATTTATGCCATGAAGAATTAAGAAA 381
 Qy 3344 GCGGACGCTTCTCAAGCGGAATCAGGTGGCACACGCTGAGGCCGAGAGGAGGATATCTCGCG 3403
 Db 382 TCTGAATGTCTAGAGAGCAGGTTGAGCATGTGAGAGCAGAGAGGAAACCTCTGCTGGT 441
 Qy 3404 GAAGCCGACAAATCACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAGGATATCTG 3463
 Db 442 GAGGTTGAAAGCCATTAATTTGGAAGCTTACTATTCTATTTTCAGGATCCCGAGTATCTG 501
 Qy 3464 TACTTCTGTAGTACTACATACAGGTGGTGTGATCTGATGCTGCTCATCAAACTGGGC 3523
 Db 502 TATCTGATTAAGAAATCTCCCGGTGGTGTATGATGACCTTGCTCATGAGGAGGAG 561
 Qy 3524 ATTTTCGAGGAGGAACTGGCCAGATTCTACATCGCGAGGTCACTGCGCCCTGGACAGC 3583
 Db 562 ACATTACGGGAGACATGTGCCAGATTTATATGCTCAAGTGTCTCGGCCATTTGAATCC 621
 Qy 3584 GTTCACAAAATGGGCTTATTCACAGAGACATCAAGCCCTGACAACATCTCATCATAGG 3643
 Db 622 ATACACAGATACAACTATATTCATAGGAGATCAAAACCTGATAACCTACTTTTGGACAA 681
 Qy 3644 GACGACACATAAAGCTCAACGACTTTGGCTGTGCACGGGATTCGATGGAGGACACAA 3703
 Db 682 GATGGCACATGAACTCTCGGACTTTGGGCTCTGTGAAGCCCTCTTGTATGTAGAAATTTA 741
 Qy 3704 TCGAAGTACTACAGGAGA-----ACGGCAATCACTCGCCGAGG 3743
 Db 742 CCTTCAATTCAGGAGAAATAGGCCACGGATGATGAACATATGTACAGAACCTATGATGTT 801
 Qy 3744 ACTCATGAGCCCTGGGAGGAAATCTCGGAGACGGACCGAACCCACCGTGTGGAGA 3803
 Db 802 GATAGATGCTTTCTCGACACTGATATAACAAGAGAGAGCTGGCGAGTCCCGCAGAA--CAAC 859
 Qy 3804 GCGACGGATGCGGATCACCAGAGTCTCGGCCACTCGCTGGTGGGCGACCCGCACT 3863
 Db 860 TTCAGCAATGGCAGATGAATCGCAGAAACTAGCATTTTCACTGTGGGAACACCGGACT 919
 Qy 3864 ACATAGTCCCGAGGTGCTGGAGAGTGGGTACACGACGTGTGCGACTTACTGGAGCG 3923
 Db 920 ATATTGCTCTGAAGTTTGTCTGAAGAAAGGATATGGCATGGAATGTGATTTGGTGGT 979
 Qy 3924 TGGCGTCTATCTTAYGAGATGCTGGTGGTGCAGCCGCCCTTTCTGGCCAAAGTCCG 3983
 Db 980 TAGTGCAATTTATGACGAAATGCTGTTGGTATCTCTCTTTTATGCTGATGACCTA 1039
 Qy 3984 TGGAAACCAAAAGGTCTCACTAAGTGGGAGAAACCTGATATTCGCGCGCAGCGCG 4043
 Db 1040 TATCAACTTGCAGAAAGATCTCCATTCATGGAGAAACCATTTGAATTTCTTCGAGGATCGGA 1099

Query Match	3.38:	Score 188.6:	DB 10:	Length 1818:	
Best local Similarity	54.98:	Preq. No. 7.4e-32:			
Matches 469:	Conservative	2:	Mismatches 356:	Indels 28:	Gaps 4:
QY	3108	AGAATCAGCTGGAGAAGGAGATGCACAAAGTGGGACATCCCGATCAGACCCCAATCGAGA	3167		
DB	729	AAAAGAAGTTAGAAAAGGTGTAAGAAGAGGCCCTAAAAGATGAGGAGAAACGACTCC	788		
QY	3168	TGAGGAAATGCTGAACCAAGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGG	3227		
DB	789	GGAGTCAGCATGCTCGGAAGGAACACAGATTTCTTCGTTTGAAGAGAACAGACTTG	848		
QY	3228	ACAAGAGCATGTTCTGTCRAAACTGAAGCCCATGGAGTGGGTGCATTTGGCGAGGTAA	3287		
DB	849	GATTGGAAGATTTTGAGTCTCTTAAAAAGTAATAGGCAGAGGAGCATTTGGTGAGGTAC	908		
QY	3288	TGTTGAGCAAAATCGNATACCTCGAACCACTTTTGATCGCATGAAAACCCCTCGGAAAGCGG	3347		
DB	909	TTGTTCCAGAGAAGAAAGATA---CGGGACATGTGTATGCAATGAAATACTCCGTAAGCAG	965		
QY	3348	ACGTTCTCAAGCGGAATCAGGTGGCACAGCTGAAGCGCGAGAGGATATCTCTCCGAAG	3407		
DB	966	ATATGCTTGAAGAAGACAGGTTGGCCACATTCGTCGGAGCGTGACATTTCTAGTGGAGG	1025		
QY	3408	CCGACAAATACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGATATCTGTACT	3467		
DB	1026	CAGACAGTTTGTGGGTGTGAATAATGTTCTATAGTTTTTCAGGTAAGCTAAACCTCTACC	1085		
QY	3468	TTGTGATGGACTACATACCAAGTGGTGTATCTGATGCGTGCATCAAACTGGGCATTT	3527		
DB	1086	TAATCATGGAGTTCYTGSCCTGGAGGGACATGATGACCTTGTGTATGATAAAAGACACT	1145		
QY	3528	TCGAGGAGGAACCTGGCCAGATTTACATCGCGAGGTCACTTCGCGCGGTGGACACGCTTC	3587		
DB	1146	T-GACGAGAGGAGACTTCAGTTTTATATAGCAGAAACAGATTTAGCCATAGACTCTAATTC	1204		
QY	3588	ACAAATGGGCTTCATTTACAGAGACATCAAGCCTGCACACATCTATCATAGGGAGC	3647		
DB	1205	ACCACTTGGATTCATCCACAGAGACATCAAAACGACACCTCTCTTTGGACACGAAG	1264		
QY	3648	GACACATAAAGCTCACCGACTTTGGCCTGTGCAGGGATTCGATGGAGCGCAACTCGA	3707		
DB	1265	GCACATGTGAACATTTCTGACATTTGGTCTTTTGACAGACTGAAAAGCACATAGGACAG	1324		
QY	3708	AGTACTACAGAGAACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAGGAT	3767		
DB	1325	AATTTTA-TAGGAATCTGAACCACAGCCTCCAGTGATTTCACTTCCAGAACATGAAT	1383		
QY	3768	ACTCCGAGACGGACCGAAGCCACCGTGTGTGGAGGCGACGGATCGCGCATCAACAA	3827		
DB	1384	TCCAAAAGGAAAGCAGAACCTGGAAGAAATAATAGCTCAG-----	1425		
QY	3828	GAGTCTCGCCCACTTCGTTGTGGGCACCCCGAATCATATAGTCCGAGGTCTCGAGA	3887		
DB	1426	-----CTAGCCCTTCCACAGTAGGCACCTCTCTGACTACATTCCTCTGAGGTGTTATGC	1480		
QY	3888	GRATGGGTACACGAGCTGTCGACTACTGGAGCGTGGCGTCACTCTTATYAGAGATGC	3947		
DB	1481	AGACCGGGTACACRAGCTCTGTGTTGGTGCCTTGGGCTCATCATGTATGAGATGC	1540		
QY	3948	TGTTGGGTACGCGC	3962		

	Query Match	2.9%;	Score 164.6;	DB 10;	Length 678;
	Best Local Similarity	63.2%;	Pred. M. 8.4e-27;		
	Matches 285;	Conservative	0;	Mismatches 161;	Indels 5; Gaps 2;
3078	TCAAAGTCGTATCGCCAGCGCACGCTATCGCAAGAATCAGCTGGAGAGAGATGCACAAG	3137			
3138	TGGGACTGCCCGATCAGACCCAAATCGAGATGAGAAATGCTGACCAAAAGGAGAGCA	3197			
63	TTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTCCCAAAAGRATCTA	122			
3198	ACTACATTCGATTGAAGCGGCCAAGATGGACAGAGCATGTTCTGCTAACTGAAGCCCA	3257			
123	ATTACATCCGCTTTAAAGGCTAAATGGCAAGTCTATGTTTGTGAAGATAAAGACAC	182			
3258	TTGGAGTGGGTGCATTTGGCGAGGTACGCTGGTGAGCAAAATCGATACCTCGAACCAT	3317			
183	TAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAGAAAAGTAGATA--CTAAGGCTT	239			
3318	TCTATTCGATGAAACCCCTCGGGAAAGCGGAGTTCTCAAGCGGAATCAGGTGGCACAG	3377			
240	TGTATGCAACAAAACCTCTCGAAGAAAGATGTCTCTTCGAAATCAAGTCGCTCATG	299			
3378	TGAAGCCGAGAGGATATCCTCGGGAAGCGGACAATACTGGGTGGTGAAGTTGTTACT	3437			
300	TTAAGGCTGAGAGAGATATCTGGCTGAAGCTGACATGAATGAATGGTAGTTCGTCATATT	359			
3438	ACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGATGGCATCAACAGGTGGTGATC	3497			
360	ATTCAATCCAAAGATAAGGCCCATTTATCCTT--GTAATGGCTACATTCCTTGGGGTGATA	417			
3498	TGATGTCGTGCTCATCAAACTGGGCATTTT	3528			

Db 418 TGAAGAGCCCATTAATTANAATGGGCATCTT 448

RESULT 9

US-09-764-868-196

; Sequence 196, Application US/09764868

; Patent No. US20020168711A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: Ptz32

; CURRENT APPLICATION NUMBER: US/09/764,868

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 196

; LENGTH: 734

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-868-196

Query Match 2.88; Score 161.2; DB 9; Length 734;

Best Local Similarity 52.68; Pred. No. 4.9e-26;

Matches 383; Conservative 4; Mismatches 317; Indels 24; Gaps 1;

Qy 3313 CCATTTGTATCGATGAACCCCTGCGAAGCGGACGCTTCTCAAGCGGAATCAGGTGGC 3372

Db 26 CCATATCTCATGATGAAGATATTGAGAAAGCTGATATGTTGAAAAGAGCAGGTGGC 85

Qy 3373 ACAGTGAAGCGCGAGGGATATCTCTCGCGAAGCGGACAAATAACTGGGTGTAAGTT 3432

Db 86 CCATATCCGAGCAGAAAGAGATATTTGGTARAAGCAGATGGTCCCTGGGTGTAAGAT 145

Qy 3433 GTACTACAGCTTCAGGACAGGATATCTCTACTTTGTGATGACTACATACAGGTGG 3492

Db 146 GTTTTACAGTTTTCAGGATAGAGAACTTTATCTAATCATGGAATTTCTCCCTGGAG 205

Qy 3493 TGATCTGATGTGCGTGCATCAAACTGGCGATTTTCAGGAGGAACTGGCCAGATTCTA 3552

Db 206 TGACATGATGACATTTGCTTAATGAGAAAGACWCCTTGACAGAGAGGAACACAGTCTTA 265

Qy 3553 CATCGCCGAGCTACCTCGCGCGTGGACAGGTTTCACAAAATGGCTTCATTCACAGAGA 3612

Db 266 CATTTTCAGAGACTGTTCTGCGAATAGATCGCATCCACAGTTGGTTTCATCCATCGGA 325

Qy 3613 CATCAAGCTCAGACATACACTCATGATAGGAGCGGACACATAAAGCTCACCAGTTGG 3672

Db 326 TATTAAGCCAGACACCTTTTATTGGATGCCAAGGTCATGTAAATTTATCTGATTTTGG 385

Qy 3673 CCTGTGACGGGATTCGATGGAGCGCACAACTCGAAGTACTACCAGGAGAACGGCAATCA 3732

Db 386 TTTATGTACGGGATTAAGAAAGCTCACAGGACTGAATTTTATAGAAATCTCACACAA 445

Qy 3733 CTCGGCCAGACTGGATGGAGCCCTGGGAGGAATATCTCCAGAGACGAGCCAGCCAC 3792

Db 446 CCCACCAAGTACTCTCATTTTCAGAACATGAACCTCAAGAGGAAAGCAGAAACTT----- 501

Qy 3793 CGTGTGGAGGGCAGCGATCGCGATCCACAGAGTCTCGCCCACTCGCTGGTGGG 3852

Db 502 -----GGAGAAGAACAGGAGCAACTGGCATATTCACAGTTGG 541

Qy 3853 CACCCGCAACTACATAGCTCCCGAGGTGCTGGAGAGAGTGGGTACAGCGAGCTGTGGGA 3912

Db 542 GACACAGATTACATTTGCTCAGAAAGTATTATCGAGACTGGTTACACAAATTTGTGGA 601

Qy 3913 CTACTGGAGCTGGCGCTATCTCTTAYAGATGCTGGTGGGTGAGCGCCCTTTCTGGC 3972

Db 602 CTGCTGGTCTTTGGGAGTATTATGTATGAAATGCTAATAGGATATCCACCTTTCTGCTC 661

Qy 3973 CAACAGTCCGTGGAACGCAACAAAGGTATCACTGCGGAGAAACSCGTGCATATTC 4032

Db 662 TGAACACCTCAAGAGACATACAGAAAAGTATGAACCTGGAAAGAAACTCTGGTATTTC 721

Qy 4033 GCCGCAGG 4040

Db 722 TCCAGAGG 729

RESULT 10

US-09-938-842A-633

; Sequence 633, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; TITLE OF INVENTION: SAME, AND METHODS OF USE

; FILE REFERENCE: SCRIPI300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866

; PRIOR FILING DATE: 2000-08-24

; PRIOR APPLICATION NUMBER: US 60/264,647

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/300,111

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 633

; LENGTH: 1398

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-633

Query Match 2.58; Score 144.6; DB 9; Length 1398;

Best Local Similarity 59.78; Pred. No. 3.1e-22;

Matches 262; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

Qy 3249 TGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAACTGGTGGAGCAAAATCGATACCT 3308

Db 410 TGAAGTTGTTGGGAAGTGGCTTTGGGAAGTCTACCGAGTGAGGAAAGGAGACT 469

Qy 3309 CGAACCATTTGATCGGATGAAACCCCTGCGGAAAGCGGAGTCTTCAAGCGGAATCAGG 3368

Db 470 C---TGAGATATACGCTATGAAGTCTATGAGAAAGATCAGATTATGAGAGAACCATG 526

Qy 3369 TGGCACAGCTCAAGCCGAGAGGATATCTCCGCGGAGCCGACAAATAACTGGGTGGA 3428

Db 527 CTGAATACATGAAGCTGAGCGCGATATCTTACAAAATTTGATCACCCTTCATTTGTTT 586

Qy 3429 AGTTGTACTACAGCTTCCAGGACAGGATAATCTGTACTTTGTGATGGACTACATACCAG 3488

Db 587 AACTTAATATCTTTTTCAGACCAAGTACAGGCTGTATCTTTGTCTCGACTTTATAAAG 646

Qy 3489 GTGGTGATCTGATGCTGCTCATCAAACTGGGCGATTTTCGAGAGAGGAATGGCCAGAT 3548

Db 647 GAGGTCTATCTTTTCTTCAACTTATCACCAGGCTTTTCAGAGAGGACTTTGGCTCGAG 706

Qy 3549 TCTACATCGCGAGGTCACTGCGCGGAGAGGCTTCAAAAATGGCTTCATTCACA 3608

Db 707 TGTACATGCAAGATCTGCTCTCGAGTTTCCCATCTCCATGAGAAAGGCAATATGCACA 766

Qy 3609 GAGACATCAAGCCCTGACAACTACTCATCGATAGGAGCGGACACATAAAGCTCCCGACT 3668

Db 767 GAGATCTCAAAACCCGAAACATACACTCATGGATACAGATGGCCACCTGATGTACAGATT 826

Qy 3669 TTGGCTGTGACCGGATTT 3687

Db 827 TTGGTTTGAAGCAAGGAATT 845

RESULT 11

US-09-938-842A-2503

; Sequence 2503, Application US/09938842A

Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2503
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2503

Query Match 2.4%; Score 138.6; DB 9; Length 1416;
Best Local Similarity 58.7%; Pred. No. 6.4e-21;
Matches 259; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 3247 ACTGAAGCCATTTGGAGTGGTGCATTTGGCGAGGTAAACCTGGTGGAGCAAAATCGATAC 3306
Db 426 ATTGAAGTTGTGGCAAGAGTGCATTTGGAAAGTGTACCGAGTGGAGGAAAAGACAC 485
QY 3307 CTCGAACCATTTGTATCGGATGAAACCCCTCGGAAAGCGGAGCTTCTCAAGCGGAATCA 3366
Db 486 GTC---TGATATACCGATGAAGTTCATGAGAAAGATAAATGTTGAGAGATCA 542
QY 3367 GGTGGCAGCGTGAAGCCGAGAGGATATCTCGCGGAAGCCGACAAATAACTGGGTGGT 3426
Db 543 TGTGATATACATGAAAGCCGAGGCGGATATCTTAACCAAAATCGATCATCTTTTCATTGT 602
QY 3427 GAAGTTGTACTACAGCTTCAGGACAGGATATCTGTACTTTGTGTGATGACTACATACC 3486
Db 603 GCAACTTAATATCTTTTCAGACCAAAATACAGATTGTATCTTGTGACTTTATAAA 662
QY 3487 AGTGGTGTATCTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGAACTGGCCAG 3546
Db 663 CGCAGGTGATCTTTTCTCCAGCTATACCAAGGCTTTTCAGGAGGACTTGGCTG 722
QY 3547 ATTCTACATCGCCGAGTCACTCGCCGCGGAGACAGGTTCAAAAATGGGCTTCATTCA 3606
Db 723 TGTGTACACTGCAGAAATCTCTCTGCAGTTTCCCATCTCCATGAGAAAGGCATAATGCA 782
QY 3607 CAGAGACATCAAGCCGTGACAACTACTCATCGATAGGCGGACACATAAAGCTCACCGA 3666
Db 783 TAGAGATCTTAACCTGAAACATACTCATGGACGTAGATGGCCATGTGTGTTAACTGA 842
QY 3667 CTTTGGCTGTGACCGGATT 3687
Db 843 TTTTGGTTAGCAAGAATT 863

RESULT 12
US-09-880-107-2146
; Sequence 2146, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2146
; LENGTH: 3061
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L07597
US-09-880-107-2146

Query Match 2.4%; Score 134.6; DB 10; Length 3061;
Best Local Similarity 58.0%; Pred. No. 7.3e-20;
Matches 258; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 3239 TTCGTCAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAAACGCTGGTGAGCAAA 3298
Db 230 TTCGAGCTCTCAAGTTCCTGGCCAGGATCCTTTGGCAAAAGTCTTCCTGGTGGCGAAA 289
QY 3299 ATCGATACCTCGAACCATTGTTGATGCG---ATGAAAACCCCTGCGAAAGCGGACGTTCTC 3355
Db 290 GTCACCCGGCTGCACAGTGGCCACCTGTATGCTATGAAGTGTCTGAAGAAGCAACGCTG 349
QY 3356 AAGCGGAATCAGTGGCAGACACGCTGAAGCCGAGAGGATATCTTCGCGGAAGCGGCAAT 3415
Db 350 AAAGTACGTGACCGCTCCGCGCAAGATGGAGAGACATCTGCTGATGATAAATCAC 409
QY 3416 AACTGGTGTGAAGTTGTACTACAGCTTCAGGACAAAGATATCTGTACTTTGTGATG 3475
Db 410 CCATCTGCTGAGCTGCACATGCTTCCACAGCCGAGGCAAGCTCTATCTCATCTG 469
QY 3476 GACTACATACAGGTGGTGTGATGCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 3535
Db 470 GACTTCTGCTGGTGGGACCTCTTCAACCCGCTCTCAAAAGAGGTGATGTTTCACGGAG 529
QY 3536 GAATCGCCAGATTTACATCGCCGAGGTACCTGGCCGTGGACAGCGTTTCACAAATG 3595
Db 530 GAGATGTGAAGTTTACCTGGCCGAGCTGGCTCTGGGCTGGATCACCCTGCACAGCTG 589
QY 3596 GGCTTCATTCACAGACATCAAGCTGCACACATATCATCATAGGCGGACGACACATA 3655
Db 590 GGTATCATTTACAGAGACCTCAAGCTGCAGAACATCTTCTGATGAGGAGGCCACATC 649
QY 3656 AAGCTCACGACTTTGGCTGTGCA 3680
Db 650 AAACCTCACTGACTTTGGCCTGAGCA 674

RESULT 13
US-09-971-845-1
; Sequence 1, Application US/09971845
; Patent No. US20020132247A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILLAINATHAN
; TITLE OF INVENTION: DYSTROPHIA MYOTONICA PROTEIN KINASE
; FILE REFERENCE: (DM-PK) AND ITS USES
; FILE REFERENCE: KINE028CON
; CURRENT APPLICATION NUMBER: US/09/971,845
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/21479
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/238,558
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3407
; TYPE: DNA

```

: CURRENT APPLICATION NUMBER: US/09/796,692
: CURRENT FILING DATE: 2001-03-01
: PRIORITY APPLICATION NUMBER: 60/186,126
: PRIORITY FILING DATE: 2000-03-01
: PRIORITY APPLICATION NUMBER: 60/190,479
: PRIORITY FILING DATE: 2000-03-17
: PRIORITY APPLICATION NUMBER: 60/200,545
: PRIORITY FILING DATE: 2000-04-27
: PRIORITY APPLICATION NUMBER: 60/200,303
: PRIORITY FILING DATE: 2000-04-28
: PRIORITY APPLICATION NUMBER: 60/200,779
: PRIORITY FILING DATE: 2000-04-28
: PRIORITY APPLICATION NUMBER: 60/200,999
: PRIORITY FILING DATE: 2000-05-01
: PRIORITY APPLICATION NUMBER: 60/202,084
: PRIORITY FILING DATE: 2000-05-04
: PRIORITY APPLICATION NUMBER: 60/206,201
: PRIORITY FILING DATE: 2000-05-22

```

1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024, 2025, 2026, 2027, 2028, 2029, 2030, 2031, 2032, 2033, 2034, 2035, 2036, 2037, 2038, 2039, 2040, 2041, 2042, 2043, 2044, 2045, 2046, 2047, 2048, 2049, 2050, 2051, 2052, 2053, 2054, 2055, 2056, 2057, 2058, 2059, 2060, 2061, 2062, 2063, 2064, 2065, 2066, 2067, 2068, 2069, 2070, 2071, 2072, 2073, 2074, 2075, 2076, 2077, 2078, 2079, 2080, 2081, 2082, 2083, 2084, 2085, 2086, 2087, 2088, 2089, 2090, 2091, 2092, 2093, 2094, 2095, 2096, 2097, 2098, 2099, 2100, 2101, 2102, 2103, 2104, 2105, 2106, 2107, 2108, 2109, 2110, 2111, 2112, 2113, 2114, 2115, 2116, 2117, 2118, 2119, 2120, 2121, 2122, 2123, 2124, 2125, 2126, 2127, 2128, 2129, 2130, 2131, 2132, 2133, 2134, 2135, 2136, 2137, 2138, 2139, 2140, 2141, 2142, 2143, 2144, 2145, 2146, 2147, 2148, 2149, 2150, 2151, 2152, 2153, 2154, 2155, 2156, 2157, 2158, 2159, 2160, 2161, 2162, 2163, 2164, 2165, 2166, 2167, 2168, 2169, 2170, 2171, 2172, 2173, 2174, 2175, 2176, 2177, 2178, 2179, 2180, 2181, 2182, 2183, 2184, 2185, 2186, 2187, 2188, 2189, 2190, 2191, 2192, 2193, 2194, 2195, 2196, 2197, 2198, 2199, 2200, 2201, 2202, 2203, 2204, 2205, 2206, 2207, 2208, 2209, 2210, 2211, 2212, 2213, 2214, 2215, 2216, 2217, 2218, 2219, 2220, 2221, 2222, 2223, 2224, 2225, 2226, 2227, 2228, 2229, 2230, 2231, 2232, 2233, 2234, 2235, 2236, 2237, 2238, 2239, 2240, 2241, 2242, 2243, 2244, 2245, 2246, 2247, 2248, 2249, 2250, 2251, 2252, 2253, 2254, 2255, 2256, 2257, 2258, 2259, 2260, 2261, 2262, 2263, 2264, 2265, 2266, 2267, 2268, 2269, 2270, 2271, 2272, 2273, 2274, 2275, 2276, 2277, 2278, 2279, 2280, 2281, 2282, 2283, 2284, 2285, 2286, 2287, 2288, 2289, 2290, 2291, 2292, 2293, 2294, 2295, 2296, 2297, 2298, 2299, 2300, 2301, 2302, 2303, 2304, 2305, 2306, 2307, 2308, 2309, 2310, 2311, 2312, 2313, 2314, 2315, 2316, 2317, 2318, 2319, 2320, 2321, 2322, 2323, 2324, 2325, 2326, 2327, 2328, 2329, 2330, 2331, 2332, 2333, 2334, 2335, 2336, 2337, 2338, 2339, 2340, 2341, 2342, 2343, 2344, 2345, 2346, 2347, 2348, 2349, 2350, 2351, 2352, 2353, 2354, 2355, 2356, 2357, 2358, 2359, 2360, 2361, 2362, 2363, 2364, 2365, 2366, 2367, 2368, 2369, 2370, 2371, 2372, 2373, 2374, 2375, 2376, 2377, 2378, 2379, 2380, 2381, 2382, 2383, 2384, 2385, 2386, 2387, 2388, 2389, 2390, 2391, 2392, 2393, 2394, 2395, 2396, 2397, 2398, 2399, 2400, 2401, 2402, 2403, 2404, 2405, 2406, 2407, 2408, 2409, 2410, 2411, 2412, 2413, 2414, 2415, 2416, 2417, 2418, 2419, 2420, 2421, 2422, 2423, 2424, 2425, 2426, 2427, 2428, 2429, 2430, 2431, 2432, 2433, 2434, 2435, 2436, 2437, 2438, 2439, 2440, 2441, 2442, 2443, 2444, 2445, 2446, 2447, 2448, 2449, 2450, 2451, 2452, 2453, 2454, 2455, 2456, 2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 26

Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880.107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2340
LENGTH: 1635
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: Genbank Accession No. US20020142981A1 M34182
US-09-880-107-2340

Query Match: 2.2%; Score 127; DB 10; Length 1635;
Best Local Similarity: 56.8%; Pred. No. 2.5e-18;
Matches 254; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 3238 GTTCGTCAAACTGAAGCCCATTTGGAGTGGGTGATTTGGCGAGGTAACGCTGGTGAGCAA 3297
DB 181 GTTCGAAACGGCTCAGGACGCTGGGCATGGGCTCCTTCGGGGGGGTGATGCTGGTGAAGCA 240
QY 3298 AATCGATACCTCGAACCATTTGTATGGATGAACCCCTGGGAAAGCGGAGCTTCTCAA 3357
DB 241 CCAGGAGACCGGGGCCA---CTACGCCATGAAGATCCTCAACAGCAGAAGGTGGTGAA 297
QY 3358 GCGGAATCAGGTGGCACACGTGAAGCGGAGGGGATATCCTCGGGAAGCCGACAAATA 3417
DB 298 GATGAAGCAGGTGGAGCACATCTGAACGAGAGCGCATCTTCGAGGCGATCGACTTTCC 357
QY 3418 CTGGGTGGTGAAGTTCTACTACAGCTTCCAGGACAGGATAATCTGTACTTTGTATGGA 3477
DB 358 GTTCCTGGTCAAGCTCCAGTTCTCTTTAAGGACAACTCTACCTGTACCTGGTGATGA 417
QY 3478 CTATATACCAGGTGGTGATCTGATGCTGCTCATCAACCTGGGCATTTTCGAGGAGGA 3537
DB 418 GTACGTGCGGGTGGGAGATGTTCTCCGCCCTACAGCGCGTGGGAAGGTTTAGCGAGCC 477
QY 3538 ACTGGCCAGATTCTACATCGCGAGGTACCTGCGCCGTGGACAGCGTTACAAAAATGGG 3597
DB 478 CCATGCGCTTTTCTATGCGCCCGAGGTCTCTGCGCGTCCAGTACCTACACTCGCTGA 537
QY 3598 CTTTATTCACAGACATCAAGCCTGACACATATCTATAGGAGCGGACACATAAA 3657
DB 538 CCTCATCCACCGGACCTGAGGCCCGGAGAACTCTCTATCGACCAGCAGGCTACCTGCA 597
QY 3658 GCTACCGACTTTGGCCTGTGCACGGG 3684
DB 598 GGTGACGGACTTCGGTTTCGCCAAGCG 624

Search completed: January 16, 2003, 22:07:03
Job time : 303.731 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 4944.5 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1 atagcagcagcgagca.....gccttgccaatttagttac 5720

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
c 1	890.8	15.6	1101	17	CNS014SB
2	643.6	11.3	662	13	BI637464
3	340.8	6.0	930	14	B0898648
4	334.8	5.9	737	14	BM950668
5	320.6	5.6	644	9	A1517339
6	304	5.3	689	13	BI331257
					GH28053.5
					BI331257 602982912

7	300.6	5.3	962	17	CNS0396F
8	291.2	5.1	590	13	BM052863
9	284.4	5.0	579	13	BI964175
10	281	4.9	957	14	BM801311
c 11	273	4.8	1006	17	CNS077GT
12	262	4.6	280	13	BI632945
13	255.2	4.5	863	10	BE620135
14	255.2	4.5	736	12	BG538495
15	255.2	4.5	824	13	BI525737
16	254	4.4	648	13	BI155504
17	251.6	4.4	948	17	CNS0257Y
18	233.6	4.1	580	13	BM174548
19	232.6	4.1	496	9	AI594372
20	231	4.0	407	9	AA592233
21	215.6	3.8	559	10	AW642458
22	209	3.7	468	12	BF774005
23	207.2	3.6	875	13	BI664597
24	206.8	3.6	425	10	BB751204
25	200.4	3.5	528	14	BM706472
26	198.4	3.5	578	10	BE093083
27	198	3.5	670	17	CNS02CSZ
28	197.6	3.5	848	9	AU121423
29	190.8	3.3	1699	11	AY104378
30	189.4	3.3	460	10	AW227865
31	186.4	3.3	1172	13	BI118231
32	184.8	3.2	442	10	AW653741
33	183.2	3.2	522	13	BI290748
c 34	182.2	3.2	813	17	CNS03JMP
35	181.8	3.2	447	9	AA023402
36	181.6	3.2	587	14	BQ769246
37	177.2	3.1	692	13	BJ300999
38	176	3.1	642	10	BB644520
39	174.2	3.0	771	12	BE866996
40	173.2	3.0	614	13	BJ121611
41	173.2	3.0	819	14	BM779602
42	171	3.0	797	13	BI109675
43	170.6	3.0	545	14	BQ143694
44	170	3.0	541	10	AW563655
45	169.8	3.0	688	9	AU177122

ALIGNMENTS

RESULT 1	CNS014SB	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	CNS014SB				
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC BACN12B11 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL104549				
VERSION	AL104549.1	GI:5616563			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster				
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk - . This Drosophila melanogaster BAC library (DrosBAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Genevieve Payan. It has been constructed in the vector pBelOAC11.				
FEATURES	Location/Qualifiers				

[illegible]

Db 241 GATTTCAGAGTACAAAATACCATAGGAATATACAGAAATATACACCTCTGCGATACACGG 300
 QY 1251 CGACCAACGGAGCGCAACGATGCACTTACTCTGACTATACACCGCCAGCAGCGATGG 1310
 Db 301 CGACCAACGGAGCGCAACGATGCACTTACTCTGACTATACACCGCCAGCAGCGATGG 360
 QY 1311 AGCCGCCACCTCCCGCTCTCTGCTCGGAGCTGGTATACCGCGCGCGCGCCGCAATG 1370
 Db 361 AGCCGCCACCTCCCGCTCTCTGCTCGGAGCTGGTATACCGCGCGCGCGCCGCAATG 420
 QY 1371 TAGGTACGCGCGAGCGCGCTCCATATCCGTATCCGGTGTGGCGCTTGAGTGTGGGTG 1430
 Db 421 TAGGTACGCGCGAGCGCGCTCCATATCCGTATCCGGTGTGGCGCTTGAGTGTGGGTG 480
 QY 1431 TGGCCAAACGGAGCTGSCCAAGATGATCAGCGCCCTATGCCAAACAACATGATCCCGA 1490
 Db 481 TGGCCAAACGGAGCTGSCCAAGATGATGACGGCCCTATGCCAAACAACATGATCCCGA 540
 QY 1491 AGCCGAGCATCAACGGGACGCGGAGCAGTCACTTACTGCGCTGCACTGCGGCTCTGG 1550
 Db 541 AGCCGAGCATCAACGGGACGCGGAGCAGTCACTTACTGCGCTGCACTGCGGCTCTGG 600
 QY 1551 ACTCGGAGCGCGGTAGTCCCGATCGGACAGCGCCCATTCGCACACACCCACCGCGGA 1610
 Db 601 ACTCGGAGCGCGGTAGTCCCGATCGGACAGCGCCCATTCGCACACACCCACCGCGGA 660
 QY 1611 GC 1612
 Db 661 GC 662

RESULT 3
 BQ898648
 LOCUS
 DEFINITION AGENCOURT_8489599 Lupski_dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6180847-5', mRNA sequence.
 ACCESSION BQ898648
 VERSION BQ898648.1 GI:22290662
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 930)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Cloning Distribution: Agencourt Bioscience Corporation
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLNL13564
 High quality sequence stop: 637.
 Location/Qualifiers
 1. 930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6180847"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
 NotI; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACCGGCTCCG-3' and

5'-GACTAGTTCTAGATCGGAGCGCGCCT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies."

BASE COUNT 266 a 229 c 261 g 174 t
 ORIGIN
 Query Match 6.0%; Score 340.8; DB 14; Length 930;
 Best Local Similarity 70.2%; Pred. No. 6.4e-50;
 Matches 486; Conservative 0; Mismatches 202; Indels 4; Gaps 2;
 QY 2993 GAGGAGGAGCGCAAGAGTTCGCGCATCAGGAGTACTCGCGCAAGCTTCAAGTTCTTC 3052
 Db 93 GACGAGAGAGAGAGAGTACGCGATCAAGAGCTACTCGCCATAGCCCTTTAAGTTCTTC 152
 QY 3053 ATGGAGCAGCATATAGAAAGCTGATCAAGTCTGTATCGCCAGCGCAGCTATCGCAAGAT 3112
 Db 153 ATGGAGCAGCATGCGGAGATGTATCAAACTTACCAGCAGAGGTTAAACCGGAGGCTG 212
 QY 3113 CAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCCAATCGAGATGAGG 3172
 Db 213 CAGCTGGAGAGAGAAATGGCCAAAGCTGGACTCTCTGAACTGACGAGGAGCAGATGCGG 272
 QY 3173 AAAATGCTGAACAAAAGAGAGCAACTACATTTGAAAGCGCGCCAAAGATGACACAAG 3232
 Db 273 AAGATCTCTTACCAGAAAGAGTCTAATTACACAGGTTAAGAGGGGCCAAGATGACACAAG 332
 QY 3233 AGCATGTTCTGTCAAACTGAAGCCCATTTGAGTGGGTGATTTGGCGAGGTACACGCTGGTG 3292
 Db 333 TCTATGTTTGTCAAGATCAAAACCTCGGATCGGTCTTTGGAGAAAGTGTGCTTCTGCT 392
 QY 3293 AGCAAAATCGATACCTCGAACCATTTGTATGGATGAAACCCCTCGGAAAGCGGACGTT 3352
 Db 393 TGTAAAGGTGGA--CACTCAGCGCCCTGTACGCCATTGAAGACCCCTAAGGAAAAAGATGTC 449
 QY 3353 CTCAGCGGGAATCAGGTGGCAGCAGCTGAAGCGGAGAGGAGATATCTCGCGAAAGCCGAC 3412
 Db 450 CTGAACCGGATCAGGTGGCGCCACGTCAGGCGGAGAGGAGATCTCTGCCGAGGACAC. 509
 QY 3413 AATAACTGGTGGTGAAGTTGTACTACAGTTCTCCAGGACAAGGATTAATCTGTACTTTGTG 3472
 Db 510 AATGATGGTGGTGTCAAACTCTACTACTCTTCCAAAGACAAGACAGCGCTGTACTTTGTG 569
 QY 3473 ATGGACTACATACCAAGTGGTGTGATCTGATGCTGCTGCTCATCAAACTGGGCAATTTCCAG 3532
 Db 570 ATGGACTACATCCCTGGTGGGAGCATGATGAGCTGCTGATCCGGATGGAGGTCTTCCCT 629
 QY 3533 GAGGAACCTGGCAGATTCTACATCGCGAGGTCACTCGCGGTGACAGCGTTTCACAAA 3592
 Db 630 GAGCACTTGGCCGGTCTTACATCGCAGAGCTGACTTTGGCCATTGAGAGTTCACCAAG 689
 QY 3593 ATGGCTTCTATTACAGAGACATCAAGCTGACACATCTCATCTATGATAGGAGCGGACAC 3652
 Db 690 ATGGCTTCTATCCACGAGACATCAAGCTGATATAATTTGATAGATCTGGATGCTC- 748
 QY 3653 ATAAAGCTCACCGACTTTGGCTGTGACGGG 3684
 Db 749 ATAAACTCACAGATTTTCGGCTCTGCACGTG 780

RESULT 4
 BQ950668
 LOCUS
 DEFINITION UI-M-EH0p-buu-b-12-0-UI.r1 NIH_BMAP_EH0p Mus musculus cDNA clone
 IMAGE:5686955 5', mRNA sequence.
 ACCESSION BQ950668
 VERSION BQ950668.1 GI:19434258
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

FEATURES
 source

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 737)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-femail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
National Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES source

Location/Qualifiers

1..737
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5686955"
/clone_lib="NIH-BMAP_EH0p"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/notes="Organ: brain; Vector: pYX-Asc; Site: 1: EcoR I;
Site: 2: Not I; The library was constructed according to
Bonatido, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGCAGC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT ORIGIN

Query Match. 5.9%; Score 334.8; DB 14; Length 737;
Best Local Similarity 69.1%; Pred. No. 7.6e-49;
Matches 505; Conservative 0; Mismatches 213; Indels 13; Gaps 3;
QY 2929 CAGCTGCAAGAAGATCAAGCAGCCCTGCGCCATCCCGGAGCGCAAGAAGATCTCCAAGGA 2988
DB 20 CAGAGACAAAAGCAGATTCAGACCTCCCGGTGCTCGCAAGAATAGC----- 71
QY 2989 GAGGAGGAGGAGCGGAGAGTTCGCATCAGCAGTACTCCCGCAGCCCTCAAGTT 3048
DB 72 -AGAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATGCTTCAAAAT 130
QY 3049 CTTTCATGGAGCAGCAGATAGAGAACGTCATCAAGTCGTATCGCCAGCGCAGTATCGCAA 3108
DB 131 CTTTCATGGAGCA-CAGCTGAGAGATGTCATCAAACTTACCAGCAAGAGTCAAGCGGAG 189
QY 3109 GAATCAGCTGGAGAGAGATGCACAAAGTGGGACTGCGCCGATCAGACCCAAATTCAGAT 3168
DB 190 GCTACAGCTGGAGCAGGAAATGCCAAAGCTGGGCTCTGFGAGCGGAGCAGCAGCAGAT 249
QY 3169 GAGGAAATGCTGAACCAAGAGAGAGACACTACATTCGATTGAAGCGGCCAAGATGGA 3228
DB 250 GAGGAGATGCTCTACCAAGAGAGAGTCTAATCAACCGCTGAGAGGCGGCAAGATGGA 309
QY 3229 CAAGAGCATCTTCGCTAAACTGAAGCCCATTTGGAGTGGGTGCTATTTGGCGAGGTAACGCT 3288
DB 310 CAAGTCCATGTTTGTGNAATCAAGACTCTAGGCATCGGTGCTTTGGGAGTGTGCCT 369

QY 3289 GGTGACGAAAATCGATACCTCGAACCACTTTGTATGCGATGAAACCCCTCGGGAAGCGGA 3348
DB 370 CGCTTTTAAGCTGGACA--CTCACGCTCTGTAGCGCATGAAGACTCTCAGGAAGAAGGA 426
QY 3349 CGTTCACACGCGAATCAGGTGGCAGACGTAAGGCGGAGAGGATATCTCCGCGAAGC 3408
DB 427 TGTCTGAACCGGAATCAAGTGGCCCATGTCAGGGCTGAGAGGACATCTCTGGCTGAAGC 486
QY 3409 CGACAATAACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGCAAGGATAATCTGTACTTT 3468
DB 487 AGCAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGGCAAGGACAGCCTGTACTTT 546
QY 3469 TGTGATGAGTACATACCAAGTGGTGGTGTGATGTGCTGCTCATCAAACTGGGCAATTT 3528
DB 547 TGTGATGAGTACATACCAAGTGGGATGATGATGAGCTCTGATCAGGATGGAGGCTTT 606
QY 3529 CGAGGAGGAACCTGCCAGATTCTACATCGCGGAGGTACCTGCGCCCTGGACACCGTTCA 3588
DB 607 CNGTGAGCACCTGGCGCGCTTCTACATTGCAGAGTTGACCTGGCCATTGAAAGTGTCCA 666
QY 3589 CAAAATGGGCTTCATTACAGACAGACATCAAGCCTGACACATCTACTCATCAGGAGCGG 3648
DB 667 CAAGATGGGCTTTATCCACCGGAGACATCAAGCCTGACACATCTACTCATCAGGATGG 726
QY 3649 ACACATAAAGC 3659
DB 727 TCATATTAAAGC 737
RESULT 5
LOCUS AI517339 644 bp mRNA linear EST 19-APR-2001
DEFINITION GH28053.5prime GH Drosophila melanogaster head pot2 Drosophila
melanogaster cDNA clone GH28053 5prime similar to U29608: wts
FBgn0011739 PID:903942 SPREMBL:Q24096, mRNA sequence.
ACCESSION AI517339 GI:4420439
VERSION AI517339.1
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 644)
AUTHORS Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G.M.
TITLE BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd., Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu
Based upon the presence of vector sequence at both ends, this
sequence has been determined to be the complete cDNA insert.
Plate: 280 row: E column: 5
High quality sequence stop: 561.
FEATURES
Location/Qualifiers
1..644
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="GH28053"
/clone_lib="GH Drosophila melanogaster head pot2"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/note="Organ: head; Vector: pot2; Site: 1: EcoRI; Site: 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. plasmid cDNA library."
BASE COUNT 145 a 162 c 159 g 178 t
ORIGIN

TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 962)

AUTHORS Genoscope

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000)

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES Location/Qualifiers

source

1..962

BASE COUNT 200 a 303 c 282 g 176 t 1 others

ORIGIN

Query Match 5.3%; Score 300.6; DB 17; Length 962;

Best Local Similarity 74.2%; Pred. No. 7.1e-43;

Matches 394; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 3133 CAAAGTGGGACTGCCGATCAGACCAATCGAGATGAGAAATGCTGAACAAAGGA 3192

Db 435 CCAGTGGGCTTTCACAGAGGCTCAGGAGCAGATGCCGATGCTGCTGCCAAGGA 494

QY 3193 GAGCACTACATTCGATTAAGCGCCCAAGATGAGCAAGAGATGCTGCTGAACCTGAA 3252

Db 495 GTCCAACTACATCGGCTGAAGCGCCCAAGATGAGCAAGAGATGCTGCTGAACCTGAA 554

QY 3253 GCCCATTTGAGTGGGTCATTTGGCGAGTAACGCTGGTGAGCAAAATCGATACCTGAA 3312

Db 555 GACCTCGCATCGGGCCCTTCGGCGAGGTGTGCTGGCCAGGA---GGAGGACACGG 611

QY 3313 CCATTTGATGCGATGAACACCTCGGAAAGCGGACGCTTCTCAAGCGGATCAGTGGC 3372

Db 612 CTCGCTCTACGCGATGAACACCTCGGCAAGAGGACGCTGCTGAGGAATCAGTGGC 671

QY 3373 ACACGTGAAGCGGAGAGGATATCTCGGGAAGCGGACAAATACCTGGTGTGAAGTT 3432

Db 672 TCACGTCAAGCGGAGAGGACATCTCGGCAAGAGCGGACAAACGAGTGGTGGTGTCT 731

QY 3433 GTACTACAGTTCACGAGCAAGGATAATCTGTACTTTGTGATGACTACATACAGTGG 3492

Db 732 CTACTACTCTTCCAGGACAGGACACCTGTACTTTGTGATGAGTACATCCAGGAG 791

QY 3493 TCATCTGATGCTGCTCATCAACTGGGATTTTCGAGGAGAACTGGCCAGATTCTA 3552

Db 792 GGACATGATGAGCTGTCTCCGCTGGGATCTTCAGGAGGAGCTGGCCAGTTCTA 851

QY 3553 CATCGCGAGGTACCTCGCGCTGGAGACGCTTCACAAATGGCTTCATTCACAGAGA 3612

Db 852 CATCGCGAGGTACCTCGCGCTGGAGACGCTTCACAAATGGCTTCATTCACAGAGA 911

QY 3613 CATCAAGCCTGACACATACATCATGATAGGACGGACACATAAAGCTCAC 3663

Db 912 CATCAAGCCTGACACATCTCATGACAGGAGGACGGACACATCAAGCTCAC 962

RESULT 8

BM052863

LOCUS

DEFINITION

1e67c07.y3 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens

CDNA clone IMAGE:5672052 5' similar to TR:Q920W4 Q920W4 LARGE TUMOR SUPPRESSOR 1; mRNA sequence.

ACCESSION

BM052863

VERSION

BM052863.1 GI:16808642

KEYWORDS

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 429.

Location/Qualifiers

1..590

BASE COUNT 164 a 138 c 165 g 123 t

ORIGIN

Query Match 5.1%; Score 291.2; DB 13; Length 590;

Best Local Similarity 69.4%; Pred. No. 3.6e-41;

Matches 411; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 3063 ACATAGAGACGTGATCAAGTCTGATCGCGCGACGCTTCAAGATCAGCTGAGCA 3122

Db 1 ACCTGGAGAAATGATCAAAACCTTACCAGCAAGGTTAAACCGAGGCTGCGAGTGGAGC 60

QY 3123 AGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAATGCTGA 3182

Db 61 AGCAATGCGCAAGCTGGAGCTGTGTGAAGCTGAGCAGGAGCAGATCGGAAGATCCTCT 120

QY 3183 ACCAAAGGAGCAACTACATTCGATTCGAAGCGGCCCAAGATGACAGAGCATGTTCCG 3242

Db 121 ACCAGAAAGAGTCTAATTAACACAGGTTAAAGAGGCCCAAGATGACAGGATCTATGTTG 180

QY 3243 TCAAACTGAGCCCATTCGATGAGTGGTGCATTTGGCGAGGTAAACGCTGTGAGCAAAATCG 3302

Db 181 TCAAGATCAAAACCTTGGGGATCGGTGCTTTTGGAGAAAGTGTGCTTCTTGTAAAGTGG 240

QY 3303 ATACCTCGAACCATTTGTATGCGATGAAACCTTCGCGAAAGCGGACCTTCTCAAGCGGA 3362

FEATURES Location/Qualifiers

1..590

BASE COUNT 164 a 138 c 165 g 123 t

ORIGIN

Query Match 5.1%; Score 291.2; DB 13; Length 590;

Best Local Similarity 69.4%; Pred. No. 3.6e-41;

Matches 411; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 3063 ACATAGAGACGTGATCAAGTCTGATCGCGCGACGCTTCAAGATCAGCTGAGCA 3122

Db 1 ACCTGGAGAAATGATCAAAACCTTACCAGCAAGGTTAAACCGAGGCTGCGAGTGGAGC 60

QY 3123 AGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAATGCTGA 3182

Db 61 AGCAATGCGCAAGCTGGAGCTGTGTGAAGCTGAGCAGGAGCAGATCGGAAGATCCTCT 120

QY 3183 ACCAAAGGAGCAACTACATTCGATTCGAAGCGGCCCAAGATGACAGAGCATGTTCCG 3242

Db 121 ACCAGAAAGAGTCTAATTAACACAGGTTAAAGAGGCCCAAGATGACAGGATCTATGTTG 180

QY 3243 TCAAACTGAGCCCATTCGATGAGTGGTGCATTTGGCGAGGTAAACGCTGTGAGCAAAATCG 3302

Db 181 TCAAGATCAAAACCTTGGGGATCGGTGCTTTTGGAGAAAGTGTGCTTCTTGTAAAGTGG 240

QY 3303 ATACCTCGAACCATTTGTATGCGATGAAACCTTCGCGAAAGCGGACCTTCTCAAGCGGA 3362

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 429.

Location/Qualifiers

1..590

BASE COUNT 164 a 138 c 165 g 123 t

ORIGIN

Query Match 5.1%; Score 291.2; DB 13; Length 590;

Best Local Similarity 69.4%; Pred. No. 3.6e-41;

Matches 411; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 3063 ACATAGAGACGTGATCAAGTCTGATCGCGCGACGCTTCAAGATCAGCTGAGCA 3122

Db 1 ACCTGGAGAAATGATCAAAACCTTACCAGCAAGGTTAAACCGAGGCTGCGAGTGGAGC 60

QY 3123 AGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAATGCTGA 3182

Db 61 AGCAATGCGCAAGCTGGAGCTGTGTGAAGCTGAGCAGGAGCAGATCGGAAGATCCTCT 120

QY 3183 ACCAAAGGAGCAACTACATTCGATTCGAAGCGGCCCAAGATGACAGAGCATGTTCCG 3242

Db 121 ACCAGAAAGAGTCTAATTAACACAGGTTAAAGAGGCCCAAGATGACAGGATCTATGTTG 180

QY 3243 TCAAACTGAGCCCATTCGATGAGTGGTGCATTTGGCGAGGTAAACGCTGTGAGCAAAATCG 3302

Db 181 TCAAGATCAAAACCTTGGGGATCGGTGCTTTTGGAGAAAGTGTGCTTCTTGTAAAGTGG 240

QY 3303 ATACCTCGAACCATTTGTATGCGATGAAACCTTCGCGAAAGCGGACCTTCTCAAGCGGA 3362

REFERENCE AUTHORS

TITLE JOURNAL

COMMENT Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center for information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 429.

Location/Qualifiers

1..590

BASE COUNT 164 a 138 c 165 g 123 t

ORIGIN

Query Match 5.1%; Score 291.2; DB 13; Length 590;

Best Local Similarity 69.4%; Pred. No. 3.6e-41;

Matches 411; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

Db 241 A---CACTCAGCCCTGTAGGCATGAACCCCTTAAGSAAAGAGGATGCTCCTGAACCGGA 297
 QY 3363 ATCAGGTGCACACGCTGAAGGCGGAGGAGGATATCTCCGGAAGCGGACATAACTGGG 3422
 Db 298 ATCAGGTGCCACGCTCAAGGCGGAGGAGGATCTCCGCGAGGCGAGCAATGAGTGGG 357
 QY 3423 TGGTGAAGTTGACTACACGCTCCAGGACAAGGATAATCTGACTTTGTGATGGACTACA 3482
 Db 358 TGGTCAAACTACTACTCTCTTCCAGACAAAGACAGCCGTACTTTGTGATGGACTACA 417
 QY 3483 TACCAGTGGTATCTGATGTCGCTCATCAAACTGGGCAATTTTCGAGGAGGAACTGG 3542
 Db 418 TCCCTGGTGGGACATGATGAGCCGTGCTGATCCGGATGGAGGTCTTCCCTGAGCACTGG 477
 QY 3543 CCAGATTCTTACATTCGCGGAGGTCACCTGCGCGTGCACAGCGTTCACAAAATGGGCTTCA 3602
 Db 478 CCGGTTCTTACATCCGAGAGCTGACTTTGGCCATTGAGAGTTCACACAGATGGGCTTCA 537
 QY 3603 TTCACAGACATCAAGCCCTGACACATACTCATCGATAGGAGCGGACACAT 3654
 Db 538 TCCACCGAGACATCAAGCCCTGATAACATTTTGATAGATCTGGATGGTCAAT 589

RESULT 9

B1964175

LOCUS

DEFINITION

B1964175 579 bp mRNA linear EST 12-MAR-2002
 ie65909.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 CDNA clone IMAGE:5671888 5' similar to TR:Q9Z0W4 Q9Z0W4 LARGE TUMOR
 SUPPRESSOR 1 ; mRNA sequence.

ACCESSION

B1964175

VERSION

EST

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 579)

Melton, D., Brown, J., Kent, G., Permut, A., Lee, C., Kaestner, K.,

Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,

Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,

Schmitt, A., Theising, B., Ritter, E., Konko, I., Bennett, J., Cardenas

, M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T.

, Jackson, Y., and Bowers, Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center For information on

obtaining a clone please contact: Juliana Brown

(brown@fas.harvard.edu) This sequence now available from the IMAGE

consortium, for clone orders contact: info@image.llnl.gov

High quality sequence stop: 432.

Location/Qualifiers

1. 579

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5671888"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;

Site_2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldi, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 161 a 136 c 163 g 119 t
 ORIGIN

Query Match

Best Local Similarity 5.0%; Score 284.4; DB 13; Length 579;

Matches 403; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 3063 ACATAGAGAACGCTGATCAAGTCTATCGCCGCGGAGGATGATCGAAGATCAAGTGGAGA 3122

Db 1 ACCTGGAGAATGCTCATCAAAACCTACCACGACAAGGTTAACCGAGGCTGCAGCTGGAGC 60

QY 3123 AGGAGATGCACAAAGTGGGACTGCCGGATCAGACCCAAATCGAGATGAGGAAAATGCTGA 3182

Db 61 AAGAAATGGCCAAAGCTGGACTCTGTGAGCTGAGCAGGAGCAGATGCGGAAGATCCTCT 120

QY 3183 ACCAAAGGAGAGCAACTACATTCGATTCAAGCGCGCCAGATGACAGAGCATGCTTCG 3242

Db 121 ACCAAGAGAGTCTATTATCAACAGGTTAAGAGGCGCAAGATGAGCAAGTCTATGTTTG 180

QY 3243 TCAAACTGAAGCCCATTTGGAGTGGGTGCTTTGGCGAGGTAACGCTGTGTGACCAAAATCG 3302

Db 181 TCAAGATCAAAACCCCTGGGATCGGTGCTTTGGAGAGTGTGCTTCTTGAAGTGG 240

QY 3303 ATACCTCGAACATTTGATGCGATGAACCCCTCGGAAAGCGGACGCTTCTCAAGCGGA 3362

Db 241 A---CACTCAGCCCTGTACGCGATGAAGCCCTTAAGAAAAAGGATGCTCTGAACCGGA 297

QY 3363 ATCAGTGGCACACCTGAAGCGCGAGAGGATATCTCCGGAAGCGGACCAATCACTGGG 3422

Db 298 ATCAGTGGCCACCTCAAGCGCGAGAGGATATCTCCGCGAGGCGACACAAATGAGTGGG 357

QY 3423 TGGTGAAGTTGACTACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGATGGACTACA 3482

Db 358 TGGTCAAACTCTACTACTCTTCCAAAGACAAGACAGCCGTACTTTGTGATGGACTACA 417

QY 3483 TACAGTGGTGTGATCTGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACTGG 3542

Db 418 TCCCTGGTGGGACATGATGAGCCTGTCTGATCCGGATGGAGGTCTTCCCTGAGCACCTGG 477

QY 3543 CCAGATTCTACATCCCGAGGTCACCTGCGCGTGCAGAGGTTTCAAAAATGGGCTTCA 3602

Db 478 CCGGTTCTTACATCCAGAGCTGACTTTGGCCATTGAGAGTGTCCCAAGATGGGCTTCA 537

QY 3603 TTCACAGACATCAAGCCTGACACATCACTCATCATGATAGG 3644

Db 538 TCCACCGAGACATCAAGCCTGATACATTTTCATAGATCTGG 579

FEATURES

source

B1964175

LOCUS

DEFINITION

AGENCOURT 6459219 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560599

5', mRNA sequence.

ACCESSION

B1964175

VERSION

B1964175.1

GI:19118134

KEYWORDS

EST

SOURCE

human.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 957)

NIH-MGC http://mgc.nci.nih.gov/.

TITLE	COMMENT
National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM12287 row: a column: 16 High quality sequence stop: 485.	

```

FEATURES
source
Location/Qualifiers
I. 957
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5560599"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: small intestine; Vector: pCMV-SPORT6;
Site.1: Not1; Site.2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT      197 a      344 c     210 g     206 t
ORIGIN
Query Match          4.9%   Score 281;   DB 14;   Length 957;
Best Local Similarity 72.3%;   Pred. No. 2e-39;
Matches 394; Conservative 0; Mismatches 145; Indels 6; Gaps
QY 3218 GCCAAGATGGACAAAGAGCTATTTCGTCAAACTGAAGCCCATTTGGAGTGCGGATTCGCAATTTGCC 3277
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 GTCAAGATGGACAAGTCTATGTTTTGTCAAGATCAAACCCCTGGGGGATCCGGTCCCTTTGGGA 62

```

Qy	3278	GAGGTAAACGCTGGT	GAGCAAAAATCAT	ACCTCGAACCAATTTG	TATGCGATGAAACCCCTG	3333
Db	63	GAAAGTGTGCTTGT	TAAAGTGGACACT	CACGCC---TGT	ACGCCATGAGAGCCCTA	119
Qy	3338	CGAAAGCGAGCGTTC	TAAGCGGAATCAGGT	TGGCACACGTGAAG	CGCGGAGAGGGATATC	3397
Db	120	AGGAAAAGGATGTCT	GAAACCGGAATCAGGT	TGGCCACGTC	AAGCGCGAGAGGGACATC	179
Qy	3398	CTCGCGAAGCCGAC	CAATAAATCTGGTGTG	AGTTGTACTACAGCTT	CCAGGACACAGGAT	3457
Db	180	CTGGCCGAGGACAG	CAATGAGTGGTGTG	CAAACTCTACTCT	CCTTCCRAGACAAAGAC	239
Qy	3458	AATCTGTACTTTGT	GATGGACTACATAC	CAGGTGGTGATCTG	ATGTCGTCTCATCAAA	3517
Db	240	AGCCTGTACTTTGT	GATGACATACATC	CTCTGGTGGGACATG	ATGACGCTCTGTATCCGG	299
Qy	3518	CTGGGCATTTTCG	AGAGAGAACTGGC	AGATCTACATCGCCG	AGGTCACCTGGCCCGTG	3577
Db	300	ATGGAGGTCTTCCT	CTGAGACCTTGGCC	CGGTCTACATTCG	CAGAGCTGACTTTGGCCATT	359
Qy	3578	GACAGCGTTTCA	AAAAATGGCTTCA	TTCACAGAGACAT	CAAGCCTGACACATCTCATC	3637
Db	360	GAGAGTGTCCAC	NAGATGGCTTCA	TCCACCCGAGACAT	CAAGCCTGATACATTTTGATA	419
Qy	3638	GATAGGACGGAC	ACATAAAGCTCAC	CGACATTTGGCCTGT	GCACGGGATTCGATGAGCG	3697
Db	420	GATCTGGATGGT	CACATTTAAACT	TCAGATTTTCGGCC	TCTGCACCTGAGGTGACT	479
Qy	3698	CACAACCTCGA	AGTACTACACAG	GAAGCGCAATCACT	CGCCGACAGGACTCGATGGAGCCC	3757
Db	480	CACAAATCCAA	TATTTATTC---	AGAAAGGGACCA	TGTGCACAGGACGACATGGAGCCC	536
Qy	3758	TGGGA	3762			
Db	537	AGCGA	541			

RESULT 11	
CNS077GT/c	
LOCUS	
DEFINITION	

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE AUTHORS

CONCLUSIONS

TITLE
JOURNAL

MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE
JOURNAL

MEDLINE
PUBMED
REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1965, Vol. 68, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Technology	1970, Vol. 1, No. 2, pp. 1-10
3. The Importance of Parental Involvement	Journal of Educational Psychology	1975, Vol. 66, No. 3, pp. 1-10
4. The Effect of Student Motivation on Learning	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 1-10
5. The Role of the School in Society	Journal of Educational Research	1985, Vol. 88, No. 5, pp. 1-10
6. The Impact of Teacher Education on Student Achievement	Journal of Educational Research	1990, Vol. 93, No. 6, pp. 1-10
7. The Importance of Teacher Evaluation	Journal of Educational Research	1995, Vol. 98, No. 7, pp. 1-10
8. The Effect of Teacher Salary on Student Achievement	Journal of Educational Research	2000, Vol. 103, No. 8, pp. 1-10
9. The Role of the School in the Community	Journal of Educational Research	2005, Vol. 108, No. 9, pp. 1-10
10. The Impact of Teacher Training on Student Achievement	Journal of Educational Research	2010, Vol. 113, No. 10, pp. 1-10

COMMENT

FEATURES
SOURC

misc_

BASE COUNT

ORIGIN

QY	3239	Matches
DB	950	

CNS077GT 1006 bp DNA linear GSS 08-JUL-2001
T7 end of clone BB0AA011G01 of library BB0AA from strain CBS 4732
of *Pichia angusta*, genomic survey sequence.

Pichia angusta, genome size sequence.
AL432675
AL432675.1 GI:12216089
GSS.
Pichia angusta.

Pichia angusta
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; *Pichia*.
1 (bases 1 to 1006)

Souciet, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
 Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
 de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
 Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S.,

Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P. and Weissenbach, J.
Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

FEB5 Lett. 487 (1), 3-12 (2000)

20584711					
11152876	(
	2	1 to 1006			
2 (bases					
plasing					
Florocto P					
Melnostru A					
Wincker D					
Artiguenave F					

Blandin, G., Llorente, B., Maipertuy, A., Wincker, P., Artiguenave, F., and Dujon, B.
Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia angusta*
Erg. Inst. 497 (1) 76-81 (2000)

FBS Lett. 487 (1), 76-81 (2000).
20584723
11152888
3 (bases 1 to 1006).
Genoscore.

Genoscope.
Direct Submission
Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces*

P. lactis var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See Table 1 for details.

keywords for description of this sequence and for the sequence of
the other extremity of this insert.

Location/Qualifiers

1..1006

Accession-Number: accnum=

```
/organism="Pichia angusta"  
/strain="CBS 4732"  
/db_xref="taxon:4905"  
/clone="BB0AA011G01"  
/cvs_lib="BB0AA01"
```

```

feature
/clone_id="BBUAA"
/note="end : T7"
complement(4. >1001)
/note="similar to Saccharomyces cerevisiae ORF YNL161w [
cark1 : strong similarity to H mavidis Hkclp protein kinase

```

Category	Count	Percentage
1 putative frameshift(s)	208	269 c
/evidence=not_experimental	272	249 t
8 others		

	208	209	212	243	0	Others
Sex						
Age						
Weight						
Length						
DB						
Score						
Pred.						
Pred. No.						
Similarity						
4.8%;						
Score 273;						
DB 17;						
Length 1006;						
59.0%;						
Pred. No. 4.9e-38;						

542; Conservative 6; Mismatches 346; Indels 25; Gaps 4;
TCCTCAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAACGCTGGTGAGCAA 3298

Est. human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 863)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/NRP/Gazdar

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9661 row: 1 column: 13
 High quality sequence stop: 706.

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3885780"
/tissue_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="PH10B (phage-resistant)"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
208 a 218 c 271 g 166 t

```

BASE COUNT
ORIGIN

Query Match	4.5%	Score 256.2	DB 10	Length 863
Best Local Similarity	75.1%	Pred. No. 4.5e-35		
Matches 334	Conservative 0	Mismatches 108	Indels 3	Gaps 1
3318	QY	TCGTATGCGATGAAACCTCGCGAAAGCGGAGCTTCTCAAGCGGAATCAGGTGGCACACG	3377	
	Db			
81	Db	TCGTAGCCCATGAAGACCCCTAGGAAAAGGATGTCCTGAACCGGATCAGGTGGCCACG	140	
3378	QY	TCGAAGCGCGAGGAGGATATCTTCGCGGAGCCGACAATAACTGGTGGTGAAGTTTGTA	3437	
	Db			
141	Db	TCGAAGCGCGAGGAGGACATCTGCGCGAGGACAGCAATGAGTGGTGGTCAAACTCTACT	200	
3438	QY	ACAGCTTCCAGGACAAGGATATCTGTACTTTGTGTGAGTACTACATACCAAGTGGTGATC	3497	
	Db			
201	Db	ACTCTCTTCCAAGACAAAGACAGCCCTGTACTTTGTGTGAGTACTACATCCCTGGTGGGGACA	260	
3498	QY	TCATGTCTGCTGCTCATCAAACTGGGCATTTTCGAGGAGGAACCTGGCCAGATTCATACATCG	3557	
	Db			
261	Db	TGATGAGCTGCTGATCCGGATGAGGTCTTCCTGAGCACCTGGCCGGGTTTCTACATCG	320	
3558	QY	CCGAGGTCCACCTGGCCCGCTGGACAGCGTTCACAAAATGGCTTCATTTCACAGAGCATCA	3617	
	Db			
321	Db	CAGAGCTGACTTTGGCCATTGAGASTGTCCAAAGATGGGCTTCATCCACCGAGACATCA	380	
3618	QY	AGCCTGACAACATCTACTCATGATAGGACCGGACACACATAAAGCTCACCGACTTTGGCGCTG	3677	
	Db			
381	Db	AGCCTGATAACATTTTGATGATCTGGATGGTCACATTAACCTACAGATTTTCGGCGCTCT	440	
3678	QY	GCACGGGATTCGGATGAGCGCACAACTCGAAGTACTACCAGGAGACGGCAATCACTCGC	3737	
	Db			
441	Db	GCATGGGTTTCAGTGGACTACAAATTCACAAATATTACC---AGAAAGGAGGCATGTCA	497	
3738	QY	CCGAGGACTCGATGGAGCCCTGGGA	3762	
	Db			
498	Db	GACAGGACGATGGAGCCCGGGA	522	

RESULT 14
BG538495

BGS538495 736 bp mRNA linear EST 03-APR-2001
 602567216F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4691955',
 mRNA sequence.
 ACCESSION BGS538495
 VERSION BGS538495.1 GI:13530728
 KEYWORDS EST,
 human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 736)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC).
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-re@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLCMI511 row: p column: 04
High quality sequence stop: 570.

FEATURES

source.

BASE COUNT
ORIGIN

Query Match

Query Match 4.58; Score 255.2; DB 12; Length 736;

Best Local Similarity 64.7%; Pred. No. 7.1e-35;
Matches 475: Conservative 3: Mismatches 236: Indels 20.

QY 3533 GAGGAAC TGGCCAGATTCTACATCGCCGAGGTCACTGCGCCGTGGACAGCGTTCACAAA 3592

Db 3 GAGCACTGGCCCGGTTCTATATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCACAAG 62

QY 3593 ATGGGCTTCATTACAGAGACATCAAGCCTGACACATACTCATCGATAGGGACGGACAC 3652

Db : 63 ATGGGCTTCATCCACCGAGACATCAAGCCTGATAACATTTTGATAGATCTGGATGGTCAC 122

QY 3653 ATAAAGCTCACC GACTTTGGCCTGTGCACGGGATTCGGATGGACGCACAACTCGAAGTAC 3712

Db
123 ATTTAACTCACAGATTTTCGGCTCTGCATGGGCTTCAGGTGGACTCACAAATCCAAATAT 182

[illegible]

WT 5715 1ACACAGG
GATCGGCAATCAC1CGCGCCAGGACTCGATGGAGGCCCTGGGAGGA-3785

DD 183 TACCAGAAAGGGAGCCATGTTCAGACAGGACAGCATGGAGCCACCGACCTCTCTGGGATGAT 242

QY 3766 -----ATACTCCGAGACGGACCGAAGCCACCGTGCTGGAGAGCGACGGATGCCGGAT 3820

D_b 243 GTGTC TAACTGTCGGTGTGGGGACAGGCTGAAGACCCCTAGAGCAGAGGGCGCGGAAGCAG 302

Qy 3821 CACCAAGAGTCCTGGCCCACTCGCTGGTGGGCACCCCGAACTACATAGCTCCGAGGTG 3880

Db 303 CACCAGAGTGCTGGCACATTCACTGGTGGGACTCCAAACTACATCGCACCCGAGGTG 362

QY 3881 CTGGAGAGRAGTGGGTACACGCAGCTGTGGCACTACTGGACGGTGGCGCTATCCTYTAY 3940

db

[illegible]

3941 GAGATGGGGGGGTCAGCCGCCCTTCTGGCCCAACAGTCCGCTGGAAACGCAACAAG 4000
 4001
 4002

DB 423 GAGATGCTGGTGGGCAGCCGCCCTTTTGGCACCTAC1CCACAGAAACCCAGCTGAAG 482

QY 4001 GTCATCAAC-TGGGAGAAAAACSCITGCATATTCCGCCCGCAGGCGCGAGTTATCCCGCGAGGC 4059

Db. 483 GTGATCAACTTGGAGAACACGGCTCCACATTCAGCCCCAGGTGAAGCTGAGCCCTGAGGC 542


```

FEATURES
source
  Location/Qualifiers
    1. 824
      /organism="Mus musculus"
      /strain="CZECH II"
      /db_xref="taxon:10090"
      /clone_ref="IMAGE:S059023"
      /clone_lib="NCI_CGAP_Lu33"
      /tissue_type="pooled lung tumors"
      /lab_host="DH10B (phage-resistant)"
      /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
      modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
      strand cDNA was prepared from mRNA obtained from pooled
      lung tumors with a Not I - oligo(dT) primer [5'
      TGTTCACCATCTCAAGTGGGCGGCGGCTCTGTTTTTTTTTTTTTTT 3'].
      Double-stranded cDNA was ligated to Eco RI adaptors
      (Pharmacia) digested with Not I and cloned into the Not
      I and Eco RI sites of the modified pT73 vector. Library
      went through one round of normalization, and was
      constructed by Bento Soares and M. Fatima Bonaldo."
      178 a 232 c 219 g 195 t
BASE COUNT
ORIGIN
Query Match          4.5%; Score 255.2; DB 13; Length 824;
Best Local Similarity 68.6%; Pred. No. 6.9e-35;
Matches 473; Conservative 2; Mismatches 190; Indels 25; Gaps 8;

```

